Computational Evidence for the Catalytic Mechanism of Caspase-7. A DFT Investigation

Gian Pietro Miscione,* Matteo Calvaresi, and Andrea Bottoni*

Dipartimento di Chimica "G. Ciamician", Universita' di Bologna, via Selmi 2, 40126 Bologna, Italy

Received: September 17, 2009; Revised Manuscript Received: February 26, 2010

We have carried out a DFT computational investigation of the catalytic mechanism of caspases, using a model system obtained from the crystallographic structure of caspase-7. In particular, we have considered the activation of the catalytic dyad (His-144 and Cys-186) and the breaking of the substrate peptide bond. We have suggested a novel mechanism for the catalytic activation, which is rather different from that usually proposed for other cysteine proteases. Following our hypothesis the activation mechanism consists of three distinct kinetic steps leading to the protonation of the catalytic His-144 and the deprotonation of Cys-186, which is activated as a nucleophile. This mechanism corresponds to a rather complex multiple proton transfer where the substrate aspartate and one water molecule act as proton shuttles. The role played by the aspartate group explains the high specificity of caspases toward substrates containing the aspartate residue that behaves as a cofactor. Apart from acting as proton shuttles and "assisting" almost all proton transfers, the two water molecules included in our model form a complex network of hydrogen bonds that involve enzyme and substrate and stabilize the charges developing on the substrate during of the reaction. We have demonstrated the existence of an alternative reaction channel leading directly from the initial complex to the peptide bond cleavage in a single kinetic step. However, this reaction pathway can be considered very unlikely since it is characterized by a high energy barrier.

1. Introduction

Apoptosis is a mechanism of programmed cell death. In contrast to a necrotic event, in apoptosis, cells play an active role in their own death in response to a variety of stimuli. Even if it appears as a cell suicide, it constitutes an essential part of life occurring during the normal development of multicellular organisms and essential for their survival. It allows the elimination of cells produced in excess, improperly developed cells, or cells that have sustained genetic damage (for instance, it destroys damaged cells such as cells infected by viruses, cells with DNA damage, cancerous cells, and cells of the immune system after they have fulfilled their function).[1−4]

In general, apoptosis is deeply involved in several human diseases characterized either by an abnormal cell growth, such as, above all, cancer, or by an excessive cell loss, such as neurodegenerative disorders (Alzheimer's and Parkinson's disease, AIDS, osteoporosis, etc). Too little apoptosis and too much apoptosis are thought to be part of the problem in the former and latter case, respectively. Therefore, the discovery of new drugs capable of blocking or enhancing apoptosis in order to block or enhance cell elimination and treat the above-mentioned disorders is now a very promising target for pharmaceutical industry and academic researchers.[5]

An intricate and specialized molecular machinery regulates apoptosis. The main component of this process is a proteolytic system involving a family of proteases known as caspases (Csp).[6−12] The word "caspases" stands for **c**ysteine-dependent **asp**artate specific prote**ases** and denotes cysteine proteases that show a strong specificity for cleaving the peptide bond following an aspartic acid residue. Since the role of caspases is fundamental in the cellular processes, modulating the effect of these enzymes is the objective of studies focusing on the apoptosis and related diseases.[13−17] Several different caspases have been identified so far and for many of them the crystallographic structure has been published.[18−26] Apart from caspase-1, -4, and -5, which are all believed to play their role as cytokine activators,[9] the majority of the members of the caspase family are usually divided in two main classes: "initiators" and "executioners".[9] The "initiator" caspases are present as inactive zymogens in the cytoplasm. After activation by a proto-apoptotic signal, they trigger downstream "executioner" caspases, which cleave substrate peptide bonds causing cell death. All caspases share an extremely high degree of homology[9,10] and this is particularly true for the two most important "executioner" caspases, i.e., caspase-3 and caspase-7.[9] These caspases have very similar specificity for substrates and inhibitors, although it has been recently reported that they are functionally distinct proteases.[27] Importantly, it has been demonstrated that the inhibition of these two caspases is sufficient for blocking apoptosis.[12] Hence, the design of inhibitors of caspase-3 and -7 can lead, in principle, to novel drugs to treat degenerative diseases.

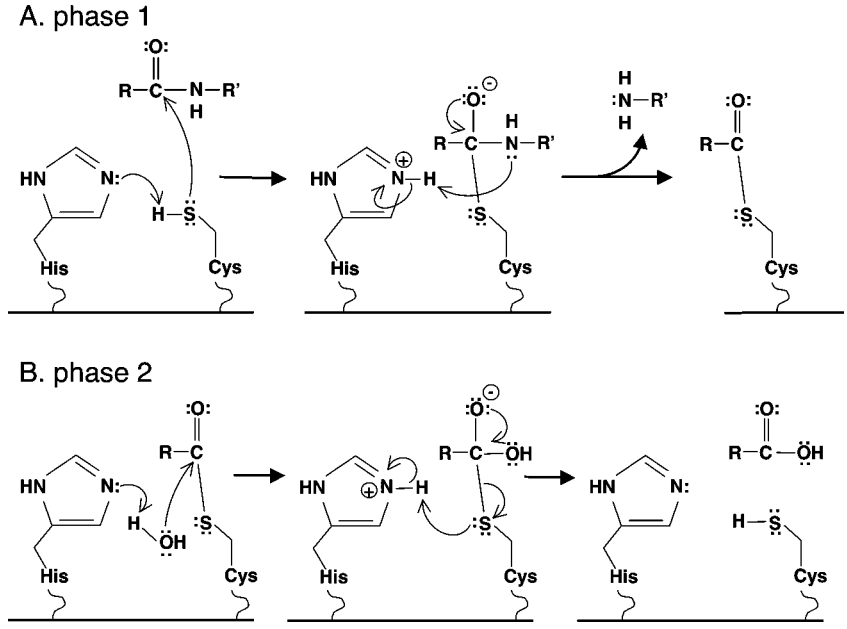
The detailed knowledge of the enzymatic reaction mechanism is essential to obtain effective inhibitors which are usually conceived as transition-state analogs. However, since these reactions are extremely fast, it is particularly difficult to determine their mechanism by means of experimental techniques. On the contrary, a computational approach represents an alternative and very successful way to obtain precious and intimate details of these very rapid processes and to confirm or reject mechanistic hypothesis.

The commonly postulated mechanism can be ideally divided in two main phases as shown in Scheme 1. In the first one, corresponding to the breaking of the peptide bond, the following steps can be recognized: (1) a histidine, behaving as a general basis, deprotonates a cysteine residue and affords a nucleophilic thiolate species (nucleophile activation); (2) attack of the

* To whom correspondence should be addressed. E-mail: gianpietro.miscione@unibo.it.

10.1021/jp908991z
Published on Web 03/12/2010

SCHEME 1: Reaction Mechanism for Cysteine Proteases



nucleophile at the carbonyl group of the aspartic peptide bond yielding a first tetrahedral intermediate; (3) protonation of the α-amino moiety by the histidine residue and formation of a potentially good leaving group; (4) cleavage of the peptide bond and consequent formation of a covalent adduct (acyl−enzyme complex).

In the second phase, the hydrolysis of the covalent adduct completes the catalytic cycle. The catalytic histidine deprotonates a water molecule and the resulting nucleophilic species attacks the carbonyl carbon of the acyl−enzyme complex, affording a new tetrahedral intermediate. In the final step, a proton moves from histidine to the sulfur atom, causing the breaking of the C−S bond, the release of the free enzyme, and the formation of the carboxy product.

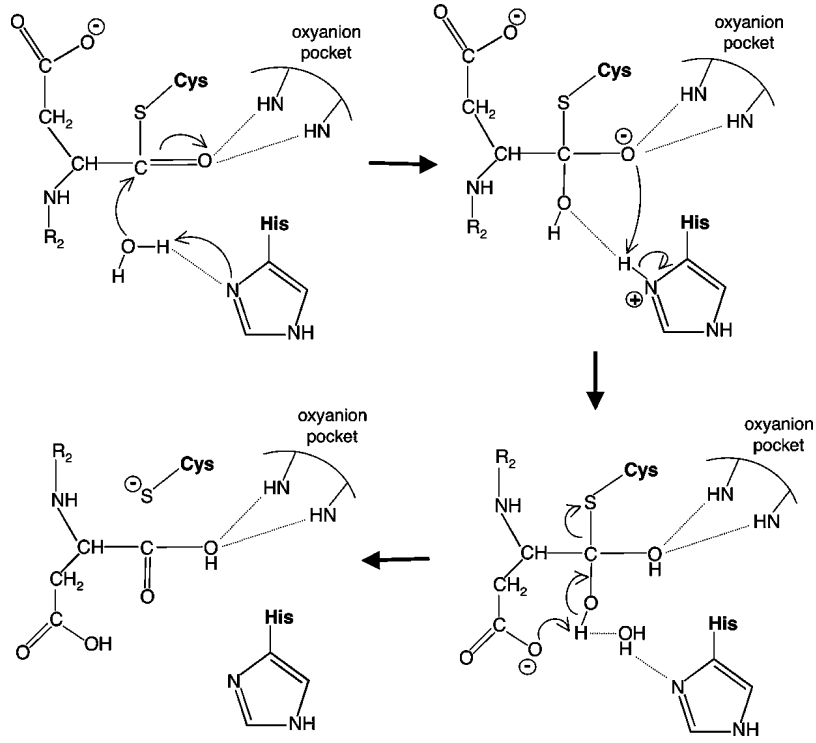
This mechanistic scheme is almost exclusively based on the presence of the typical catalytic dyad of the cysteine proteases (i.e., cysteine and histidine) and is a straightforward application of the mechanistic hypothesis commonly accepted for these kinds of proteases. However, it is not really supported by any theoretical or experimental evidence. Many structural and mutagenesis studies provide information on the caspase−substrate binding and clearly point to cysteine as the active nucleophile within the active-site,[28] even if detailed mechanistic information is rather poor at the moment.

Two main interactions are recognized to play a key role in the substrate-binding process.[28] One arises from the hydrogen bonds between the substrate carbonyl oxygen and the backbone N−H groups of the oxyanion hole, a pocket formed by a glycine residue and the catalytic cysteine (Gly-145 and Cys-186 in the caspase-7 sequence). This interaction is believed to be important in polarizing the carbonyl carbon of the scissile bond and stabilizing the subsequent charged intermediate. The other interaction involves the substrate negatively charged aspartate and the Arg-87, Arg-233, and Gln-184 residues (caspase-7 sequence). This interaction should be rather strong and can be reasonably considered as the most important factor (though not the only one) determining the extraordinary specificity of the caspases toward aspartate containing substrates.

However, some important issues concerning the mechanism proposed in Scheme 1, still remain unclear. First of all, the large histidine/cysteine distance (6−7 Å) found in caspases (larger than in normal proteases) makes very unlikely a direct proton transfer between these two residues.[8,9,28] This structural feature also argues against a general assumption stating that the Cys−His dyad exists as an ion-pair, where the proton of the cysteine residue is transferred to the histidine nitrogen atom before substrate binding. This hypothesis is wrongly based on evidence concerning only one cysteine protease (i.e., papain), which shows remarkable differences in the relative position of these two residues. Furthermore, MD calculations[29] have ruled out the ion-pair hypothesis, proving that this arrangement is rather unstable. Thus, it is reasonable to believe that the catalytic cysteine is not prepolarized[28−31] and that the deprotonation occurs during the reaction, in agreement with the pH optimum values of all caspases lying in the narrow range 6.8−7.4. As observed by Brady and co-workers,[30] the histidine residue is not even in a good position to protonate the leaving amino group, as usually supposed. These authors instead suggest that a water molecule interacting with Gly-145 (caspase-7 sequence) could more easily behave as a proton donor. The same authors emphasize an "alternative" role for histidine, postulating that it might stabilize the charge developing on the leaving group during the displacement reaction. Once protonated, this residue could interact with the ketone or thioemiketal oxygen, after the breaking of the peptide bond and the formation of a covalent bond between cysteine and Cα.

The most important mechanistic study on this subject is certainly a QM/MM investigation by Sulpizi et al.,[31] who examined in detail the reaction mechanism of caspase-3, focusing their attention on the second phase of Scheme 1. The model system used by these authors includes two water molecules. One molecule has been placed in such a way that it forms effective hydrogen bonds with the histidine residue and the substrate aspartate. The second one enters the active site from the solvent bulk during the MD simulation. The results of this investigation, which are not in agreement with the commonly postulated mechanism, can be summarized as follows (see Scheme 2): (a) histidine deprotonates a water molecule, thus enhancing its nucleophilic character; then, the "activated" water molecule attacks the acyl-enzyme complex; (b) a proton

SCHEME 2: Mechanism of Caspase-3 (Second Phase) As Proposed by Sulpizi et al.[31]



moves from histidine to the negative acyl(carbonyl) oxygen, affording a neutral gem-diol intermediate; a second water molecule from the solvent interacts with histidine and one of the diol OH groups; (c) a proton is transferred from one of the two hydroxyl groups of the diol intermediate to the aspartate side-chain carboxylate, thus determining the breaking of the C−S bond. The computations show that for the attack of the water molecule a free energy barrier of about 19 ± 4 kcal mol^{-1} must be overcome, in satisfactory agreement with the experimental results. This hypothesis suggests that the substrate might play an active role in catalysis and that the extraordinary specificity of caspases for this residue might be due not only to the strong interactions with the enzyme, but also to a direct involvement in the reaction mechanism.

In spite of these recent results, it is evident from the above discussion that the caspase mechanism is still far from being completely understood and several fundamental aspects, especially concerning the first phase of the process, remain unclear.

In the present paper, we describe the results of an accurate DFT quantum mechanical computational study on the mechanism of caspase-7. In particular, we focus our attention on the first phase of the process. We avoid a further investigation of the second phase of the catalytic cycle (see Schemes 1B and 2) since it has been recently discussed in details by other authors.[31] As far as we know, this is the first example of a complete and exhaustive investigation on this subject.

We have chosen the downstream caspase-7 because it is one of the best known and structurally characterized "executioner" caspases (together with caspase-3). In particular, the crystal structure of caspase-7 bound to the tetrapeptide aldehyde inhibitor acetyl-Asp-Val-Ala-Asp-CHO (Ac-DEVD-CHO) is available.[23] Although the model system used here is inevitably small compared to the real situation (and also to the system that could be employed using a QM/MM approach), we believe that a DFT investigation with an accurate basis set can provide detailed and important information on an enzymatic reaction, which takes place in a rather restricted spot, such as the enzyme active site. In particular, we focus our attention on the following issues: (i) the cysteine activation; (ii) the nucleophilic attack of cysteine on the carbonylic system, the consequent breaking of the peptide bond and the release of the amino-moiety; (iii) the role of water molecules within the active site; (iv) the role of the substrate aspartate residue; (v) the role of the catalytic histidine.

2. Computational Details and Choice of the Model

Our model system emulates the caspase-7 active site and is based on the crystallographic structure corresponding to the Protein Data Bank code 1FIJ (resolution = 2.35 Å).[23] In this structure, caspase-7 is covalently bound to the tetrapeptide aldehyde inhibitor Ac-DEVD-CHO. This inhibitor contains the Asp-Glu-Val-Asp (DEVD) sequence that constitutes the target of caspase-7. This complex is similar to the hypothesized acyl-enzyme intermediate characterized by a C−S bond between the substrate Asp-4′ carbonyl carbon and the sulfur atom of the catalytically active cysteine. Thus, its structure can be used to assemble a reliable model system to investigate the catalytic mechanism. To avoid confusion, it is important to underline that the numbering of the various residues that will be reported from now on follows the caspase-7 sequence and thus can be different from that given in previous papers where the sequence of other structurally characterized caspases (mainly caspase-1 or caspase-3) has been used.

All residues surrounding the substrate and taking part in the reaction, have been included in the model system which is depicted in Figure 1. These are Ser-231, His-144, Gly-145, Cys-186, Gln-184, Arg-87, and Arg-233. In order to reduce the size of the model, the portions of these residues not directly involved in the process have been replaced by smaller groups. This approximation is evident in the two pictures of Figure 1 and in Figure S1 of the Supporting Information. A point concerning Figure 1A (and Figure S1) should be stressed: since we have

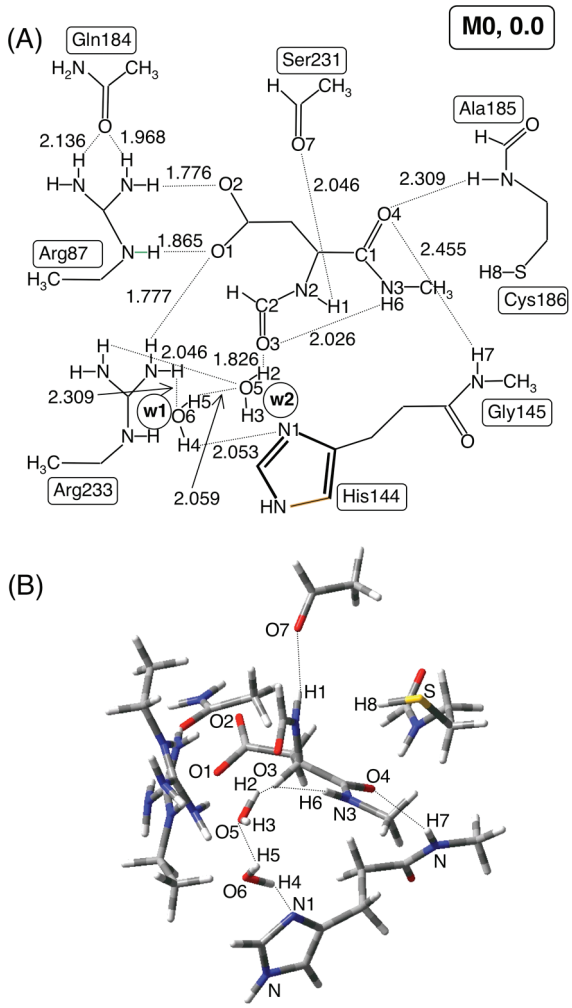


Figure 1. Two-dimensional (A) and three-dimensional (B) schematic representation of the structure of the starting complex **M0** (bond lengths are in Å).

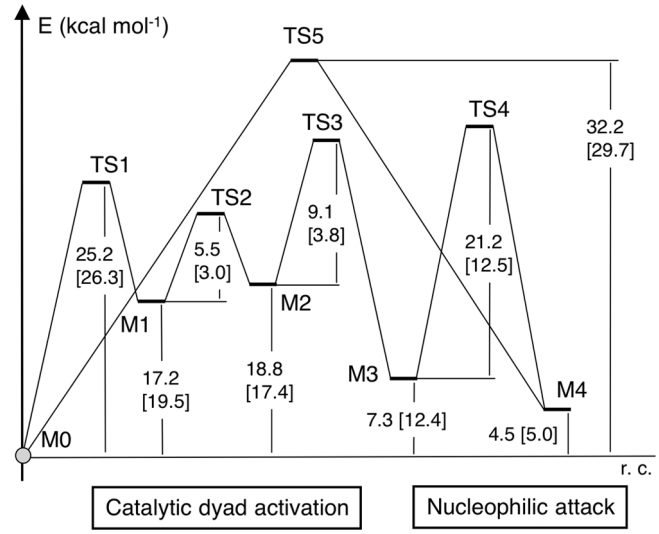


Figure 2. Energy profile computed for the first phase of the caspase catalytic cycle (catalytic dyad activation and nucleophilic attack).

adopted a two-dimensional representation, several atomic distances are not realistic and appear much longer (or shorter) than in the real protein. A more realistic three-dimensional picture of the model system is given in Figure 1B.

Furthermore, with the purpose of maintaining the geometry of the active-site cavity and, thus, emulating the partially constraining effect of the protein environment, during the geometry optimization procedure we have kept the position of the atoms of the backbone "frozen" at the crystallographic coordinates. These frozen atoms are clearly indicated in Figure S1 of the Supporting Information. Two water molecules **w1** and **w2** have been added to the model system. These water molecules have been initially positioned on the basis of the results of the MD simulations carried out by Sulpizi and co-workers[31] and form hydrogen bonds with the histidine nitrogen N1 and the carbonyl oxygen O3. Then, the initial positions have been optimized to obtain the most favorable assembling of the hydrogen bond network. Only the first two residues of the substrate (the acyl−aldehyde inhibitor) have been considered, i.e., aspartate and valine (whose size has been reduced). Since the crystallographic structure resembles the covalent adduct (acyl−enzyme complex) obtained at the end of the first phase of the process (see Scheme 1 and starting point in Scheme 2), some structural changes are needed to obtain a general model system suitable to investigate the enzymatic process from the beginning. To this purpose we have removed the C−S bond detected in the crystallographic structure and we have added a hydrogen atom to saturate the sulfur atom. Thus, the catalytic cysteine (Cys-186) is assumed to exist in its neutral state, similarly to His-144. As outlined in the introduction, this choice is consistent with several experimental data available in literature[28−31] and also with preliminary calculations that we have carried out before starting the mechanistic investigation. These computations show that a system involving a negative Cys-186 and a positive His-144 is much less stable than the neutral situation. Furthermore, the Asp-4′ carbon atom has been saturated by a $-NHCH_3$ fragment with the purpose of generating a minimal-size substrate for the enzyme and simulating the peptide bond that breaks during the first reaction phase. In the model system obtained from these approximations, the substrate is at the center of a cavity that includes seven residues and emulates the catalytic pocket of the real enzyme.

All reported DFT computations have been carried out with the Gaussian 03[32] series of programs using the B3LYP functional[33] and the DZVP basis set on all atoms. This basis is a local spin density (LSD)-optimized basis set of double-ζ quality,[34] which has been demonstrated to provide reliable results for similar systems. Single-point computations using the polarized continuous model (PCM)[35] method available in Gaussian 03 have been performed on the gas-phase optimized structures to roughly evaluate the effect of the protein environment. A dielectric constant $\varepsilon = 4.335$ has been used. This value has been recently suggested in the literature for computational studies where the PCM approach has been employed in enzymatic models involving hydrogen bonds, proton transfer and bond-breaking/bond-formation processes.[36−38]

3. Results and Discussion

In this section we discuss the results obtained in the investigation of the singlet potential energy surface for the first phase of the catalytic process. In particular we illustrate the activation of the catalytic dyad and the cleavage of the peptide bond affording the release of an amino fragment (simulating one fraction of the protein chain) and the formation of the acyl-enzyme intermediate. The corresponding energy profile is displayed in Figure 2, while a schematic representation of the structure of the various critical points is given in Figures 1 (starting Michaelis complex) and 3−5. A more detailed

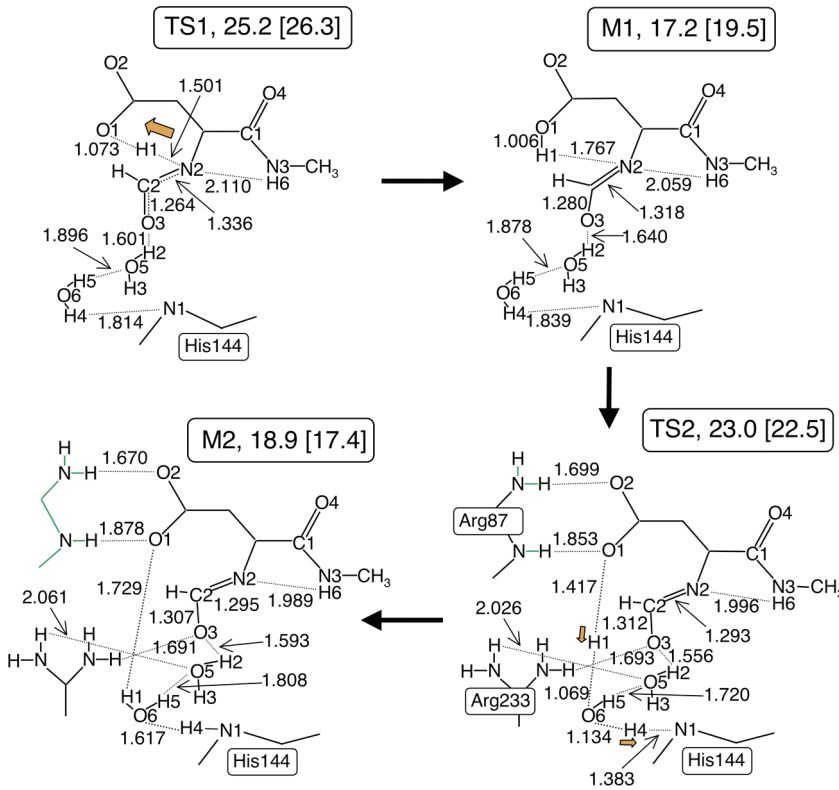


Figure 3. Two-dimensional schematic representation of the structure of the critical points **TS1**, **M1**, **TS2**, and **M2** (bond lengths are in Å). Energy values (kcal mol^{-1}) are relative to **M0** (PCM values in brackets).

representation of the various critical points is given in Figures S1−S10 of the Supporting Information.

A. Activation of the Catalytic Dyad. In the initial substrate−enzyme complex **M0** (Michaelis complex) the most important enzyme−substrate interactions involve five residues: Arg-233 and Arg-87; Cys-186 and Gly-145 forming the oxyanion hole; and Ser-231. This interaction scheme is almost completely consistent with the experimental data discussed in the introduction. The significant differences are represented by the lack of interactions between the substrate and Gln-184 (interacting with Arg-87) and the presence of a strong hydrogen bond between the Ser-231 carbonylic oxygen and the N−H group of the substrate aspartate (O7···H1 = 2.046 Å). The strongest interactions keeping the substrate anchored to the enzyme active site are those involving the aspartate carboxylate and three NH bonds of the two arginine residues, as indicated by the short NH···O distances: (Arg-87)NH···O1 = 1.865 Å, (Arg-87)NH···O2 = 1.776 Å, and (Arg-233)NH···O1 = 1.777 Å. This clearly shows the importance of aspartate in the substrate−enzyme binding process. Two hydrogen bonds (not very strong) involving Cys-186 and Gly-145 bind the carboxyl group of the peptidic bond to be broken to the oxyanion hole. These interactions enhance the electrophilic character of the carbonyl carbon C1 (the atom where the cysteine attack will take place) as indicated by the lengthening of the C1−O4 double bond (1.242 Å). Two additional strong hydrogen bonds can be recognized for the second substrate carboxyl group (C2O3 belonging to the valine residue). One involves the water molecule **w2** (O3···H2 = 1.826 Å), whereas the other is an intramolecular interaction occurring within the substrate (N3H6···O3 distance = 2.026 Å). A complex network of hydrogen bonds involving Arg-233, the C2O3 carboxyl group, and the His-144 residue binds the two water molecules **w1** and **w2** to both substrate and enzyme. In addition to the previously described C2O3···H2 interactions, other hydrogen bonds have been detected: O6H4···N1 (H4···N1 distance = 2.053 Å), (Arg-233)NH···O5 (H···O5 distance = 2.046 Å), and (Arg-233)NH···O6 (H···O6 distance = 2.309 Å). Also, a **w1**−**w2** hydrogen bond O6H5···O5 is evident (H5···O5 = 2.059 Å). It is interesting to note that, in agreement with the experimental evidence, the distance between the histidine nitrogen (N1) and the cysteine hydrogen (H8) is 7.189 Å. Thus, it is evident that the two atoms are too far away to allow a direct proton transfer between the two residues, and an alternative mechanism leading to histidine deprotonation and cysteine activation must take place. The overall picture provided by the **M0** structure corresponds to a substrate molecule surrounded and interacting with several residues. These interactions mainly involve the aspartate carboxylate. However, none of the two centers of the catalytic dyad involved in the proteolytic event (the cysteine SH group of the Cys and the hystidine nitrogen) are closely interacting with the substrate.

The first reaction step is crucial and emphasizes the role played by the substrate aspartate (carboxylate group). In the corresponding transition state **TS1** (see Figure 3) a proton (H1) moves from the substrate nitrogen N2 to the oxygen O1 of the aspartate carboxylate group. The effect of this proton transfer is a negative charge delocalized on N2, C2 and especially O3. As a consequence we observe a shortening of the N2−C2 bond (partial double bond character with distance 1.336 Å) and a lengthening of the C2−O3 bond (1.264 Å). In the transition structure, the proton transfer to O1 is almost completed, the H1···O1 distance being 1.073 Å. Interestingly, in the transition state, the interaction of the hydrogen H6 with O3 is replaced by a new hydrogen bond with the partially negative nitrogen N2 (N2···H6 = 2.110 Å). Simultaneously, a strengthening of the O3···H2 interaction (O3···H2 = 1.601 Å) is observed. This new structural arrangement allows a more effective stabilization of the developing negative charge on N2 and O3. The proton transfer from N2 to O1 requires a rotation of the aspartate

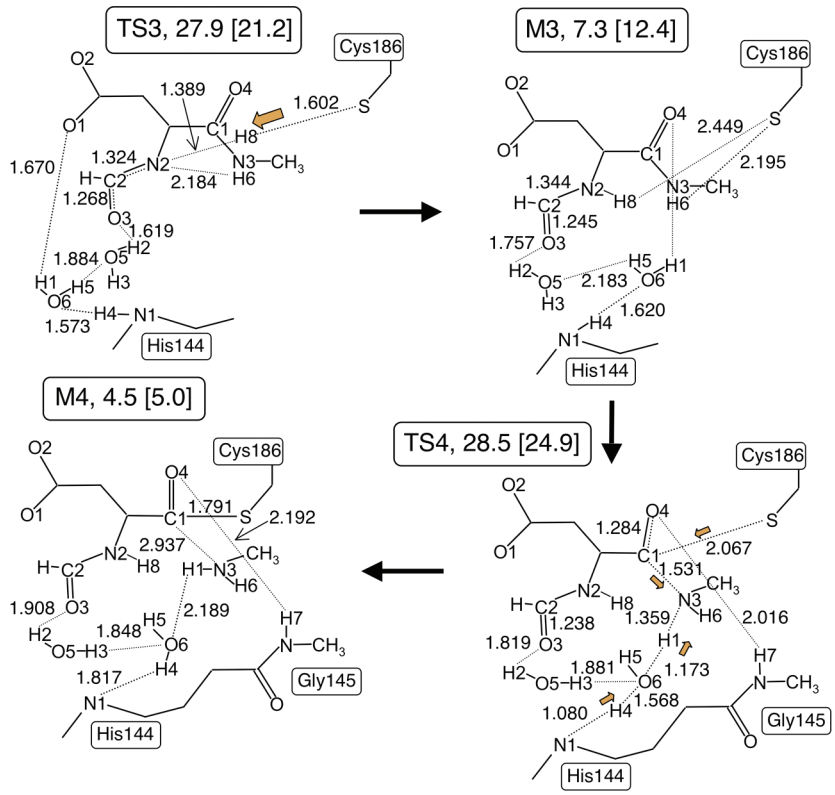


Figure 4. Two-dimensional schematic representation of the structure of the critical point **TS3**, **M3**, **TS4**, and **M4** (bond lengths are in Å). Energy values (kcal mol^{-1}) are relative to **M0** (PCM values in brackets).

carboxylate group, which causes the breaking of two of the three strong hydrogen bonds with the arginine residues. It is worth to outline that the loss of these interactions certainly contributes to the significant energy barrier (25.2 kcal mol^{-1}).

The following intermediate **M1** (see Figure 3) lies 17.2 kcal mol^{-1} above reactants (**M0**) and, in accordance with the Hammond postulate, is structurally very similar to the previously discussed transition state: the proton has been definitively transferred to the aspartate oxygen O1 and N2 forms a stable hydrogen bond with H6 (N2···H6 = 2.059 Å). It is important to note that, so far, the two water molecules and the catalytic dyad (Cys186 and His144) have not played any active role in the reaction. In particular, the network of hydrogen bonds involving **w1**, **w2**, the substrate, and His-144 has been conserved during the **M0** → **TS1** → **M1** transformation.

The analysis of the subsequent reaction step (**M1** → **TS2** → **M2**) provides further evidence of the crucial role of aspartate. In **TS2** (see Figure 3) the aspartate carboxylate group again rotates, but in the opposite direction with respect to that occurring in the previous step. This motion places H1 close enough to the O6 oxygen (**w1**) to allow a double proton transfer: one proton (H1) moves from the aspartate oxygen O1 to O6 (**w1**) and a second proton (H4) from O6 to the N2 nitrogen of histidine. The barrier for this transformation is not very high being only 5.5 kcal mol^{-1}.

TS2 leads to the intermediate **M2** (18.8 kcal mol^{-1} above **M0**) where the substrate aspartate group is again deprotonated and negatively charged as in the starting reactants. As a consequence it can form again the two initial strong hydrogen bonds with the NH groups of Arg-87. A second negative charge is delocalized on the N2, C2, and O3 centers. A significant portion of this charge is on O3, as it is evident from the large C2−O3 distance (1.307 Å) that demonstrates the C2O3 single bond character. The double negative charge on the substrate molecule, even distributed on rather distant centers, explains to a certain extent the high energy of **M2** with respect to **M0**. In the course of the **M0** → **M2** transformation the electronic situation of one water molecule (**w2**) does not change. The interactions of the **w2** oxygen O5 with one N−H group of Arg-233 and the H5 hydrogen of **w1** remain approximately constant. Similarly, the O3···H2(**w2**) interaction is maintained along the whole transformation. On the contrary, at the end of the transformation, **w1**, after having received a proton, interacts with the O1 aspartate carboxylate oxygen (O6H1···O1 = 1.729 Å in **M2**), while the hydrogen bond with Arg-233 disappears. It is important to outline that the most important result of the **M0** → **M2** process is the final protonation of the catalytic His-144. This is the result of a rather complex multiple proton transfer where the substrate aspartate and one water molecule (**w1**) act as proton shuttles.

The following reaction step (**M2** → **TS3** → **M3**) involves the catalytic cysteine residue (Cys-186). In **TS3** (see Figure 4) the sulfur proton (H8) moves to the substrate N2 nitrogen by overcoming a barrier of 9.0 kcal mol^{-1}. As a consequence, a negative charge is transferred from the substrate to the cysteine sulfur atom (which is activated as a nucleophile), the N2···H6 distance increases (from 1.989 in **M2** to 2.184 Å in **TS3**) while the C2−O3 bond length decreases (from 1.307 in **M2** to 1.268 and 1.245 Å in **TS3** and **M3**, respectively) since its double bond character is augmented. This proton transfer restores the initial charge situation on the substrate and the most important difference between the resulting intermediate **M3** (7.3 kcal mol^{-1} above **M0**) and the initial complex is the protonation state of the catalytic dyad. Thus, the final effect of the whole **M0** → **M3** process is the transfer of a proton from Cys-186 to His-144. This final arrangement i.e a negative thiolate residue, which can act as an effective nucleophile and a protonated histidine available for donating a proton, corresponds to the configuration

commonly assumed in protease mechanism as that anticipating the cleavage of a peptide bond. The most important difference with respect to the usually proposed mechanistic scheme (see Scheme 1) is given by the proton transfer mechanism. We have shown that, because of the large distance between the two catalytic residues, this occurs in three distinct kinetic steps involving a water molecule and the carboxylate group of the substrate aspartate as proton shuttles. This finding also explains the high specificity of caspases toward aspartate containing substrates.

Other interesting structural features are evident in **M3** (see Figure 4). A weak interaction between the cysteine thiolate group and the proton now bonded to N2 is still evident (H8···*S* = 2.449 Å). The sulfur atom forms another hydrogen bond with the N−H group of the peptide bond to be cleaved (H6···*S* = 2.195 Å). The adjacent carbonyl group interacts with one hydrogen atom (H1) of the water molecule **w2**, which in turn forms a hydrogen bond with **w1** (O6H5···O5 = 2.183 Å) and the N1H4 bond of histidine (N1H4···O6 = 1.620 Å). Interestingly, the histidine N1H4 bond, the water molecule **w2** and the substrate group -NHMe are all rather close and in a suitable arrangement to make possible a further proton transfer to the nitrogen of the -NHMe moiety, needed to create a good leaving group.

B. Cleavage of the Peptide Bond and Formation of the Acyl−Enzyme Intermediate. Several transformations simultaneously occur in the transition state **TS4** (see Figure 3): (i) attack of the nucleophilic thiolate on the carbonylic carbon C1 to form a new sulfur−carbon bond. The effect is an incipient tetrahedral configuration of C1 with a significant lengthening of the CO carbonylic bond (C1−O4 = 1.284 Å); (ii) cleavage of the peptide bond C1−N3: the N3−C1 bond length becomes 1.531 Å (1.335 Å in the previous minimum); (iii) protonation of the leaving group through a double proton transfer that starts from the histidine nitrogen N1 and involves the water molecule **w2**: H4 is transferred from N1 to O6, and simultaneously H1 moves from O6 to N3. Thus, the histidine residue is actually "assisting" the breaking of the peptide bond via the protonation of the leaving group. This assistance requires the presence of a water molecule, acting as proton shuttle, as suggested by Brady and co-workers.[30] The carbonyl is still involved in the same interactions with the cysteine and glycine H−N groups (oxianyon hole) observed in the starting complex. This sequence of transformations results in a rather high activation energy (21.2 kcal mol^{-1}) for **TS4**, which is 28.5 kcal mol^{-1} above the starting reactants.

The resulting intermediate **M4** (see Figure 4) is an acyl−enzyme complex where the peptide bond is definitely broken (the C1···N3 distance is 2.937 Å) and the H_2NCH_3 fragment (simulating a part of the substrate protein chain) is partially anchored to the enzyme by a hydrogen bond involving the water molecule **w2**. The two molecules **w2** and **w1** in turn participate to a complex network of hydrogen bonds with the neighboring residues.

To obtain a complete mechanistic scenario, we have explored an alternative reaction channel, which could lead in a single kinetic step from the initial reactant complex **M0** to the peptide bond cleavage (transition state **TS5** depicted in Figure 5). Within this mechanistic scheme the cysteine residue immediately attacks the carbonyl carbon C1 (S−C1 = 2.852 Å) causing the cleavage of the C1−N3 bond (1.499 Å). The two structural changes which are essential for this process (i.e., the deprotonation of cysteine to enhance its nucleophilic character and the protonation of the amino group to make it a good leaving group) are simultaneous

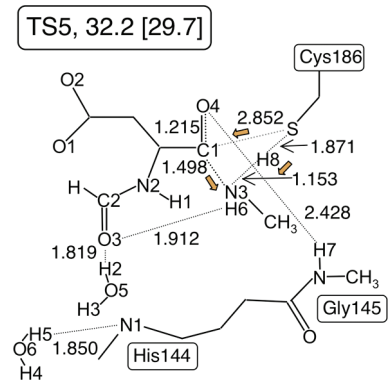


Figure 5. Two-dimensional schematic representation of the structure of the critical point **TS5** (bond lengths are in Å). Energy values (kcal mol^{-1}) are relative to **M0** (PCM values in brackets).

to the nucleophilic attack and occur through a direct proton transfer from the cysteine sulfur S to the N3 nitrogen. The synchronized displacement of the amino-group is helped by a stabilizing hydrogen bond involving the O3 oxygen, which becomes stronger with respect to the reactant situation (N3···H6 distance = 2.026 and 1.912 Å in **M0** and **TS5**, respectively). The histidine residue does not play any direct role within this simple one-step mechanism. However, the barrier associated with this process is quite high (32.2 kcal mol^{-1}) and **TS5** is 3.7 kcal mol^{-1} higher than the highest in energy critical point (**TS4**) found along the previously discussed four-step pathway. This high barrier, which is certainly caused by the geometrical distortions required in the one-step process, rules out the alternative catalytic pathway represented by transition state **TS5**.

C. Effect of the Protein Environment. The energy and activation energy values obtained in the presence of a simulated protein environment (PCM computations) are reported in brackets in Figure 2. The energies of the two intermediates **M1** and **M2** involved in the **M0** → **M3** transformation (activation of the catalytic dyad) do not change significantly (from 17.2 to 19.5 kcal mol^{-1} and from 17.4 to 18.8 kcal mol^{-1} for **M1** and **M2**, respectively). A similar trend is observed for **TS1** and **TS2**: the activation barriers for these two transition states become 26.3 and 3.0 kcal mol^{-1}, respectively (25.2 and 5.5 kcal mol^{-1} in gas-phase). A more important energy change is observed for **TS3**. In this case, the resulting activation barrier becomes much lower (3.8 kcal mol^{-1}) when compared to the gas-phase value (9.1 kcal mol^{-1}). The final effect for the catalytic dyad activation process is that, in the presence of the protein environment, the highest in energy transition state becomes **TS1**, which is 26.3 kcal mol^{-1} above the starting complex **M0**. However, this result does not change significantly the basic energetic features of the **M0** → **M3** transformation, since the energy of **TS1** from PCM computations is only slightly lower than that of **TS3** in gas-phase (27.9 kcal mol^{-1} above **M0**). A non-negligible energy variation features the **M3** → **M4** process (nucleophilic attack). In this case the transition state **TS4** is 24.9 kcal mol^{-1} above **M0** (28.5 kcal mol^{-1} in gas-phase) and the corresponding barrier becomes 12.4 kcal mol^{-1}. Thus, in the presence of protein, the energy required by the catalytic dyad activation seems to become larger than that needed by the nucleophilic attack. Finally, the alternative one-step reaction pathway represented by **TS5** remains higher in energy and disfavored with respect to the multistep channel.

In conclusion the mechanistic scenario is not remarkably affected by the protein environment. The most interesting effect suggested by our computations is a general decrease of the activation barriers when the whole protein is taken into account.

4. Conclusions

In this paper we have carried out a detailed computational investigation at the DFT level of the catalytic mechanism of caspases. In particular, we have focused our attention on the first part of the catalytic cycle (Scheme 1) where the catalytic dyad (His-144 and Cys-186) is activated and the substrate peptide bond is broken. The most relevant results can be summarized as follows:

(i) Our computations have elucidated the nature of the interactions that anchor the substrate to the enzyme active site. These mainly involve the aspartate carboxylate group and three NH bonds of the two neighboring arginine residues (Arg-87 and Arg-233). This evidence clearly shows the importance of aspartate in the substrate-enzyme recognition and binding process.

(ii) We have demonstrated the existence of a quite novel mechanism that can activate the catalytic dyad. This is rather different from that usually proposed for other cysteine proteases. That mechanism cannot occur in caspases because the histidine nitrogen and the cysteine hydrogen are too far away to allow a direct proton transfer. Our new proposed activation mechanism consists of three distinct kinetic steps. The first two steps lead to the protonation of the catalytic His-144, while in the third step Cys-186 is activated (deprotonation) as a nucleophile. The final enzyme activation is the result of a rather complex multiple proton transfer where the substrate aspartate and one water molecule act as proton shuttles. The direct participation of the substrate aspartate group to the catalytic activation (substrate-assisted catalysis) explains the high specificity of caspases toward substrates containing the aspartate residue that behaves as a cofactor.

(iii) The protonated histidine residue "assists" the breaking of the peptide bond by protonating the leaving group. The protonation involves a water molecule that again behaves as a proton shuttle. This process is made possible by the favorable arrangement reached by the histidine N−H bond, a water molecule, and the −NHMe substrate group (simulating the leaving group) after cysteine activation.

(iv) The two water molecules included in the model system play an essential role. First of all, acting as proton shuttles, they "assist" almost all proton transfers. Second, they form a complex network of hydrogen bonds involving enzyme and substrate, that allow the stabilization of the charges developing on the substrate during the course of the reaction.

(v) We have explored an alternative reaction channel leading from the initial reactant complex **M0** to the peptide bond cleavage in one single kinetic step. Within this mechanistic scheme (where the histidine residue does not play any direct role) the cysteine deprotonation, the nucleophilic attack of cysteine on the substrate carbonyl carbon, the protonation of the leaving group and its expulsion, occur simultaneously. However, the strong geometrical distortions required in this one-step reaction channel determine a quite high energy barrier (32.2 kcal mol^{-1}), which makes this alternative pathway quite unlikely.

(vi) The mechanistic picture is not remarkably affected by the inclusion of the protein environment effects simulated by PCM computations. The multistep reaction channel remains favored with respect to the one-step reaction path where activation of the catalytic dyad and breaking of the peptide bond occur simultaneously. It is worth to outline that the PCM computations suggest that the activation barriers should decrease in the presence of the protein environment.

(vii) The computed barriers (25−26 kcal mol^{-1}) are compatible, even if overestimated, with the experimental measurements carried out on other caspases (for instance, caspase-1 or caspase-3) as described in refs 21 and 31, where a free energy activation value of 17.7 kcal mol^{-1} is reported. This overestimation is not surprising if we consider the small size of the model system used here and the lacking of the protein environment only partially described by the PCM computations. However we stress once again that the aim of this paper is not a perfect estimate of the experimental activation barriers, but to employ accurate QM methods to envisage reliable mechanistic schemes.

(viii) Since we suggest that the protonated histidine residue "assists" the breaking of the peptide bond by protonating the leaving group, to evaluate the reliability of this hypothesis, it would be interesting to carry out some mutagenesis experiments where His-144 is replaced by a different residue which cannot play the role of proton acceptor/proton donor evidenced in our mechanistic scheme. Such a mutation should lead to a significant decrease of the enzyme activity. Similarly, other mutations could be carried out on Arg-87 and Arg-233 that anchor the substrate to the enzyme active site via a strong interaction with the aspartate carboxylate group.

Supporting Information Available: Figure S1. Two-dimensional schematic representation of the structure of the critical point **M0** including frozen atoms specification. Figure S2. Two-dimensional schematic representation of the structure of the critical point **TS1**. Figure S3. Two-dimensional schematic representation of the structure of the critical point **M1**. Figure S4. Two-dimensional schematic representation of the structure of the critical point **TS2**. Figure S5. Two-dimensional schematic representation of the structure of the critical point **M2**. Figure S6. Two-dimensional schematic representation of the structure of the critical point **TS3**. Figure S7. Two-dimensional schematic representation of the structure of the critical point **M3**. Figure S8. Two-dimensional schematic representation of the structure of the critical point TS4. Figure S9. Two-dimensional schematic representation of the structure of the critical point M4. Figure S10. Two-dimensional schematic representation of the structure of the critical point **TS5**. Table S1. Cartesian coordinates (Å) and total energies (ET, a.u.) for the various critical points located along the catalytic reaction pathway. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Thomson, C. B. *Science* **1995**, *267*, 1456–1462.
(2) Hengartner, M. O.; Horvitz, H. R. *Cell* **1994**, *76*, 665–676.
(3) Zheng, T. S.; Flavell, R. A. *Nat. Biotechnol.* **2000**, *18*, 717–718.
(4) Evan, G.; Littlewood, T. *Science* **1998**, *281*, 1317–1322.
(5) Reed, J. C.; Tomaselli, K. J. *Curr. Opin. Biotechnol.* **2000**, *11*, 586–592.
(6) Thornberry, N. A.; Lazebnick, Y. *Science* **1998**, *281*, 1312–1316.
(7) Stennicke, H. R.; Salvesen, G. S. *Biochim. Biophys. Acta* **1998**, *1387*, 17–31.
(8) Stennicke, H. R.; Salvesen, G. S. *Cell Death Differ.* **1999**, *6*, 1054–1059.
(9) Denault, J. B.; Salvesen, G. S. *Chem. Rev.* **2002**, *102*, 4489–4499.
(10) Chereau, D.; Kodandapani, L.; Tomaselli, K. J.; Spada, A. P.; Wu, J. C. *Biochemistry* **2003**, *42*, 4151–4160.
(11) Lee, D. *J. Biol. Chem.* **2000**, *275*, 16007–16014.
(12) Fan, T. J.; Han, L. H.; Cong, R. S.; Liang, J. *Acta Biochim. Biophys. Sin.* **2005**, *37*, 719–727.
(13) Feinstein-Rotkopf, Y.; Arama, E. *Apoptosis* **2009**, *14*, 980–995.
(14) Odonkor, C. A.; Achilefu, S. *Cancer Invest.* **2009**, *27*, 417–429.
(15) Lamkanfi, M.; Moreira, L. O.; Makena, P.; Spierings, D. C. J.; Boyd, K.; Murray, P. J.; Green, D. R.; Kanneganti, T. *Blood* **2009**, *113*, 2742–2745.
(16) Rohn, T. T.; Head, E. *Int. J. Clin. Exp. Pathol.* **2009**, *2*, 108–118.
(17) Rohn, T. T.; Head, E. *Rev. Neurosci.* **2008**, *19*, 383–393.

(18) Walker, N. P. C.; Talanian, R. V.; Brady, K. D.; Dang, L. C.; Bump, N. J.; Ferenz, C. R.; Franklin, S.; Ghayur, T.; Hackett, M. C.; Hammill, L. D.; Herzog, L.; Hugunin, M.; Houy, W.; Mankovich, J. A.; McGuiness, L.; Orlewicz, E.; Paskind, M.; Pratt, C. A.; Reis, P.; Summani, A.; Terranova, M.; Welch, J. P.; Xiong, L.; Möller, A.; Tracey, D. E.; Kamenb, R.; Wongb, W. W. *Cell* **1994**, *78*, 343–352.

(19) Wilson, K. P.; Black, J. A.; Thomson, J. A.; Kim, E. E.; Griffith, J. P.; Navia, M. A.; Murcko, M. A.; Chambers, S. P.; Aldape, R. A.; Raybuck, S. A.; Livingston, D. J. *Nature* **1994**, *370*, 270–275.

(20) Rotonda, J.; Nicholson, D. W.; Fazil, K. M.; Gallant, M.; Gareau, Y.; Labelle, M.; Peterson, E. P.; Rasper, D. M.; Tuel, R.; Vaillancourt, J. P.; Thornberry, N. A.; Becher, J. W. *Nat. Struct. Biol.* **1996**, *3*, 619–625.

(21) Mittl, P. R.; Di Marco, S.; Krebs, J. F.; Bai, X.; Karanewsky, D. S.; Priestle, J. P.; Tomaselli, K. J.; Grutter, M. G. *J. Biol. Chem.* **1997**, *272*, 6539–6547.

(22) Riedl, S. J.; Renatus, M.; Schwarzenbacher, R.; Zhou, Q.; Sun, S.; Fesik, S. W.; Liddington, R. C.; Salvesen, G. S. *Cell* **2001**, *104*, 791–800.

(23) Wei, Y.; Fox, T.; Chambers, S. P.; Sintchak, J.; Coll, J. T.; Golec, J. M.; Swenson, L.; Wilson, K. P.; Charifson, P. S. *Chem. Biol.* **2000**, *7*, 423–432.

(24) Chai, J.; Shiozaki, E.; Srinivasula, S. M.; Wu, Q.; Dataa, P.; Alnemri, E. S.; Yigong Shi, Y. *Cell* **2001**, *104*, 769–780.

(25) Blanchard, H.; Kodandapani, L.; Mittl, P. R. E.; Di Marco, S.; Krebs, J. F.; Wu, J. C.; Tomaselli, K. J.; Grütter, M. G. *Structure* **1999**, *27*, 1125–1133.

(26) Watt, W.; Koeplinger, K. A.; Mildner, A. M.; Heinrikson, R. L.; Tomasselli, G.; Watenpaugh, K. D. *Structure* **1999**, *27*, 1135–1143.

(27) Walsh, J. G.; Cullen, S. P.; Sheridan, C.; Luthi, A. U.; Gerner, C.; Martin, S. J. *Proc. Natl. Acad. Sci.* **2008**, *105*, 12815–12819.

(28) Fuentes-Prior, P.; Salvensen, G. S. *Biochem. J.* **2004**, *384*, 201–232.

(29) Sulpizi, M.; Rothlisberger, U.; Carloni, P. *Biophys. J.* **2003**, *84*, 2207–2215.

(30) Brady, K. D.; Giegel, D. A.; Grinnel, C.; Lunney, E.; Talanian, R. V.; Wong, W.; Walker, N. *Bioorg. Med. Chem.* **2003**, *7*, 621–631.

(31) Sulpizi, M.; Laio, A.; VandeVondele, J.; Cattaneo, A.; Rothlisberger, U.; Carloni, P. *Proteins: Struct., Funct. Genet.* **2003**, *52*, 212–224.

(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(33) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(34) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560–571.

(35) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117–129. Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239–245.

(36) Prabhakar, R; Morokuma, K; Musaev, D. G. *Biochemistry* **2006**, *45*, 6967–6977.

(37) Liu, H; Robinet, J. J; Ananvoranich, S; Gauld, J. W. *J. Phys. Chem. B* **2007**, *111*, 439–445.

(38) Borowski, T; Siegbahn, P. E. M. *J. Am. Chem. Soc.* **2006**, *128*, 12941–12953.

JP908991Z