Methodology article

A novel method for evaluation and screening of caspase inhibitory peptides by the amino acid positional fitness score

Atsushi Yoshimori¹, Ryoko Takasawa² and Sei-ichi Tanuma*¹,²,³

Address: ¹Research Institute for Biological Sciences, Tokyo University of Science, 2669 Yamazaki Noda, Chiba 278-0022, Japan, ²Genome and Drug Research Center, Tokyo University of Science, 2641 Yamazaki Noda, Chiba 278-8510, Japan and ³Department of Biochemistry, Faculty of Pharmaceutical Sciences, Tokyo University of Science, 2641 Yamazaki Noda, Chiba 278-8510, Japan

Email: Atsushi Yoshimori - yosimori@rs.noda.tus.ac.jp; Ryoko Takasawa - takasawa@rs.noda.tus.ac.jp; Sei-ichi Tanuma* - tanuma@rs.noda.tus.ac.jp

* Corresponding author

Abstract

Background: Since caspases are key executioners of apoptosis in cases of severe diseases including neurodegenerative disorders such as Alzheimer’s disease and Huntington’s disease, and viral infection diseases such as AIDS and hepatitis, potent and specific inhibitors of caspases have clinical potential. A series of peptide inhibitors has been designed based on cleavage sites of substrate proteins. However, these peptides are not necessarily the most potent to each caspase. Moreover, so far, it has proved to be difficult to design potent and specific peptide inhibitors of each caspase from sequence data of known cleavage sites in substrate proteins. We have attempted to develop a computational screening system for rapid selection of potent and specific peptide inhibitors from a comprehensive peptide library.

Results: We developed a new method for rapid evaluation and screening of peptide inhibitors based on Amino acid Positional Fitness (APF) score. By using this score, all known peptide inhibitors of each caspases-3,-7,-8, and -9 were rapidly selected in their enriched libraries. In this libraries, there were good correlations between predicted binding affinities of the known peptide inhibitors and their experimental Ki values. Furthermore, a novel potent peptide inhibitor, Ac-DNLD-CHO, for caspase-3 was able to be designed by this method. To our knowledge, DNLD is a first reported caspase-3 inhibitory peptide identified by using the computational screening strategy.

Conclusion: Our new method for rapid screening of peptide inhibitors using APF score is an efficient strategy to select potent and specific peptide inhibitors from a comprehensive peptide library. Thus, the APF method has the potential to become a valuable approach for the discovery of the most effective peptide inhibitors. Moreover, it is anticipated that these peptide inhibitors can serve as leads for further drug design and optimization of small molecular inhibitors.

Background

Caspases are cysteine aspartyl proteases that play critical roles during the execution of apoptosis [1-5]. Caspases have been found in organisms ranging from C. elegans to humans. The 14 types of caspase (named caspase-1 to caspase-14) identified so far in mammals are suggested to play distinct roles in apoptosis and inflammation. The main caspase cascades are catalyzed by caspases-8 and -9.
(initiator caspases) and caspases-3 and -7 (executioner caspases) [4,5]. The disregulation of caspase cascade activation is suggested to be involved as a key factor for the development of a variety of diseases, including neurodegenerative disorders and cancers [6-9].

Caspases share similarities in amino acid sequence, structure and substrate specificity, and are characterized by an almost absolute specificity in the active sites for aspartic acid in the P1 position of substrate proteins. Each active site contains a positively charged S1 subsite that binds the negatively charged P1 aspartate on the substrate [10-13]. This S1 binding site is highly conserved; therefore, all caspases cleave solely after aspartate residues. Recognition of at least four amino acids (P1-P4) in the cleavage sites is also a necessary requirement for efficient catalysis. Individual caspases have been shown to have structural differences in the predicted S2-S4 substrate binding sites which vary significantly, resulting in varied substrate specificities for the P2-P4 positions, despite an absolute requirement for aspartate in the P1 position [10-16].

The interactions between caspases and their target substrate proteins are highly flexible, so that many caspase inhibitory peptides, which are derived from native substrates, are recognized by several members of the caspase family. For instance, the sequence DEVD within poly(ADP-ribose) polymerase (PARP) [17] is known to be recognized and cleaved by caspase-3, and it has been applied to creating the tetrapeptide inhibitor Ac-DEVD-CHO. However, Ac-DEVD-CHO also shows similar inhibitory activity toward caspases-7, and -8 as well as caspase-3 [18]. Therefore, it is difficult to design selective caspase inhibitory peptides based on the sequence of known cleavage sites.

To identify the sequence with the highest binding affinity for each caspase, the X-X-X-D (X any of the 20 amino acids) peptide sequence motif is utilized for Structure-Based Virtual Screening (SBVS), which involves 205, i.e., 8000, ΔGbind evaluations. In order to show each amino acid position clearly, the X-X-D motif is expressed as P1-P2-P3-P4-D. In SBVS using a library consisting of hundreds of thousands small molecules, it is necessary to limit the computing time per compound to at most 1 min. However, even side chains of a peptide with only 4 amino acids have a maximum of 20 rotational bonds. In flexible docking using AutoDock3.0 [19] with default parameters, a computing time of about 20 min per 4-amino acid peptide on a COMPAQ Alphastation DS20E (double 833 MHz processors and 1024 MB of memory) is required. In the case of the design of a 5-amino acid peptide, it is necessary to carry out 3.2 million ΔGbind calculations, so that comprehensive peptide screening is practically difficult. Moreover, the computing time for ΔGbind calculation can be traded off against accuracy. Therefore, the most problematic aspect of ΔGbind calculation is time-consumption for SBVS.

Alternatively, various pattern-recognition techniques have been used to analyze peptide cleavage data [20,21]. A statistical technique has been used to calculate binding free energy in the analysis of resistance-evading, peptide mimetic inhibitors of HIV-1 protease [22]. These types of methods are enough to make potent peptide screening tractable. However, many experimental data points are needed to construct a highly precise scoring function.

In the present study, we have developed a new computational screening strategy based on Amino Acid Positional Fitness (APF) score for rapid evaluation of peptides interacting specifically with a target protein molecule. In our system, 1 to 10% of peptides are extracted at random from a comprehensive peptide library. The ΔGbind between the extracted peptides and the target protein are calculated using AutoDock3.0, and then the data are transferred to obtain the APF score for each amino acid position (for example; P1, P2, and P3 in caspases) using a statistical technique. Consequently, this score is able to evaluate rapidly the ΔGbind of all peptides in the comprehensive peptide library. We show here the usefulness of this APF method for the screening of potent and specific peptide inhibitors of main caspase family members.

Results and Discussion
Evaluation of binding modes and affinities of caspase inhibitory peptides by AutoDock3.0
One of goals of this work is the development of a simple and fast scoring function (APF score matrix) for the prediction of binding affinities of peptides against target protein molecules. The APF score matrix is derived from the predicted binding modes and affinities based on a standard docking program, AutoDock3.0. It is, therefore, important to verify the docking results of AutoDock3.0. Initially, we used Root of Mean Squared Deviation (RMSD) value as an indicator of the quality of the predicted binding modes. The RMSD was determined for the atomic positions between the predicted binding mode of peptide and native one as found in the crystal structure. The crystal structures of caspase-3 (1PAU), caspase-7 (1F1J), caspase-8 (1QTN), and caspase-9 (1JXQ) were used to validate the quality of predicted binding modes.

AutoDock3.0 uses a rapid grid-based lookup method for energy evaluation. The use of interaction energy grids allows the use of a sophisticated empirical free energy force field, while greatly reducing the computational demands of the docking simulation. In this study, the Lamarckian Genetic Algorithm (LGA), which combines a genetic algorithm and a local search algorithm, was used.
as the search method. Since the docking parameters of AutoDock3.0 affect the qualities of predicted binding modes and affinities, the number of energy evaluations is an important docking parameter [23]. So, we investigated the effect of the number of energy evaluations on the quality of the predicted binding mode of peptide (Figure 1). Other docking parameters were set to the default values. In the case of a fixed main chain, the RMSD values were roughly constant and less than 2 Å in all cases. In contrast, when all of rotational torsion angles were flexible (full flexible), the number of energy evaluations does have an influence. Below 1.5 million energy evaluations, the RMSD values of caspase-7/DEVD were more than 2.0 Å (Figure 1b). More than 3 million energy evaluations are required to predict the binding mode of peptide with an RMSD of less than 2.0 Å with respect to the crystal structures of the four caspases.

In Structure-Based Virtual Screening (SBVS), the computing time for evaluation of binding affinity is an important factor, as well as computational accuracy. Figure 2 shows the relationship between the number of energy evaluations and the average of the computing time on COMPAQ Alphastation DS20E. For example, in 0.7 million energy evaluations, the computing time was less than 1 minute. However, above 1.5 million energy evaluations, the computing time increased dramatically.
Fig. 2  
Effects of the number of energy evaluations on the computing time. The computing times were measured in the various number of energy evaluations. During the docking, the main chain angles of peptides are fixed (◆) or flexible (■).

Table 1: Inhibition of caspases by peptide aldehydes [18]

| Aldehyde | Caspase-3 | Caspase-7 | Caspase-8 | Caspase-9 |
|----------|-----------|-----------|-----------|-----------|
| DEVD     | 0.23      | 1.6       | 0.92      | 60        |
| IETD     | 195       | 3280      | 1.05      | 108       |
| WEHD     | 1960      | >10,000   | 21.1      | 508       |
| AEVD     | 42        | 425       | 1.6       | 48        |
| YVAD     | >10,000   | >10,000   | 352       | 970       |

Ki (nM)
evaluations and with a fixed main chain, the average computing time was about 10 min (Figure 2). This is reasonable time per peptide for constructing the APF score matrix. However, fixation of the main chain during docking may affect the quality of docking results of inhibitory peptides with arbitrary amino acid sequences. So, we compared our docking results of well-known peptide inhibitors of caspases-3, -7, -8, and -9 with their experimental data (Table 1) previously reported (Table 2). In this experiment, to generate the initial conformations of these peptides, the crystallographic structures were used for the main chain coordinate of the inhibitory peptides, and then the side chains were added. Moreover, in order to exclude van der Waals contacts of side chains, the TINKER software package [24] was used and energy minimization using the Amber94 forcefield was performed. During the minimization, the main chain atoms were fixed and RMSD gradient limits were 0.01. Kollman atom charges were assigned and the rotational bonds of the peptides were defined with AutoTors [19]. Thus, the docking parameters and conditions are determined as follows: the number of energy evaluations is 0.7 million, other parameters are default values, and main chain angles are fixed. To verify the accuracy of the docking results, we analyzed correlation coefficients between ΔG_{bind} (predicted) and log Ki values (observed). Peptide inhibitors with Ki values of >10,000 nM were set up the value of 10,000 nM. As shown in Table 2, the correlation coefficients were 0.75 (caspase-3), 0.83 (caspase-7), 0.95 (caspase-8), and 0.96 (caspase-9). The results show good correlations between predicted ΔG_{bind} values by AutoDock3.0 and experimental Ki values. These observation imply that our docking parameters and conditions (0.7 million energy evaluations and fixed main chain) are reasonable for predicting binding modes and affinities of caspase inhibitory peptides.

### Construction of APF score matrix

As a first attempt, APF score matrices for caspases-3, -7, -8, and -9 were constructed. The APF score matrix shows the P_3-P_4 preferences of a caspase (Figure 3), and if a high correlation coefficient is observed between the APF score derived from APF score matrix and ΔG_{bind}, the APF score enables the rapid prediction of ΔG_{bind} between a caspase and arbitrary peptides. Figure 4 shows the scatter plot of APF score vs ΔG_{bind} in the analysis library and prediction library. The APF score of caspase-3 showed a preference for P_4: D, V > Q, P_3: E > L > P and P_2: V > E > P (Figure 3a). The results obtained for caspase-3 indicate that the preferred recognition motif for this enzyme is (D/V)EVD. Interestingly, DEVD is a cleavage site within PARP [17] and is also the optimal recognition sequence identified from a positional scanning synthetic combinatorial library (PS-SCL) [14]. The correlation coefficient for the caspase-3 analysis library was -0.71 (Figure 4a). To verify the predictability of APF score, the correlation analysis was performed using a prediction library comprising of 40 peptides. The correlation coefficient was -0.64 (Figure 4a).

Since the APF score matrix of caspase-7 showed a preference for P_4: D > E > C, P_3: F > E > D, P_2: P > I, S > L, the preferred recognition motif for this enzyme is D(F/E)VD. DEVD, known as the optimal recognition sequence, was not shown to be a sequence with the highest APF score from the APF score matrix. However, a good correlation coefficient between APF score and ΔG_{bind} was confirmed (Figure 3b). Although known optimal recognition sequences for caspase-8 and caspase-9 were not shown from the APF score matrix to be sequences with the highest APF score, the correlation coefficients between APF score and ΔG_{bind} in the analysis library were -0.68 and -0.76, respectively (Figure 4c and 4d).

### Table 2: Prediction of binding affinities of caspase inhibitory peptides by AutoDock3.0

|                | Caspase-3 | Caspase-7 | Caspase-8 | Caspase-9 |
|----------------|-----------|-----------|-----------|-----------|
| **DEVD**       | ΔG_{bind} (kcal/mol) | -16.83    | -17.09    | -16.91    | -15.79    |
|                | RMSD (Å)  | 0.795     | 1.306     | 0.632     | 0.727     |
| **IETD**       | ΔG_{bind} (kcal/mol) | -16.07    | -15.27    | -16.85    | -15.72    |
|                | RMSD (Å)  | 0.85      | 1.188     | 0.486     | 0.729     |
| **WEHD**       | ΔG_{bind} (kcal/mol) | -15.38    | -12.34    | -14.67    | -14.63    |
|                | RMSD (Å)  | 0.837     | 3.358     | 1.6       | 1.279     |
| **AEVD**       | ΔG_{bind} (kcal/mol) | -15.82    | -15.72    | -16.1     | -15.59    |
|                | RMSD (Å)  | 0.84      | 1.591     | 0.31      | 0.771     |
| **YVAD**       | ΔG_{bind} (kcal/mol) | -11.2     | -11.39    | -14.18    | -14.24    |
|                | RMSD (Å)  | 2.201     | 2.674     | 1.493     | 0.912     |
| **Corr. Coeff. of ΔG_{bind} and log Ki (R)** | 0.75 | 0.83 | 0.95 | 0.96 |
Figure 3

APF score matrices of caspases. The APF score matrices indicate the interaction preferences of amino acids with (a) caspase-3, (b) caspase-7, (c) caspase-8, and (d) caspase-9 at each position (P4-P2). If an APF score is greater than 0, the amino acid tends to be favored at a given position. $\Delta G_{th}$ (threshold value of $\Delta G_{bind}$) of caspases-3, -7, -8, and -9 are -11.05, -10.69, -11.03, and -12.12 kcal/mol, respectively.
The scatter of APF score vs $\Delta G_{\text{bind}}$ for the Analysis library and Prediction library. The solid line is the result of a linear regression fit, and $r$ denotes correlation coefficient. (a) caspase-3, (b) caspase-7, (c) caspase-8, and (d) caspase-9.
Consequently, it is revealed that the APF score matrices of all caspases have the ability to predict $\Delta G_{\text{bind}}$, although known optimal recognition sequences for caspases-7, -8, and -9 were not shown directly.

**Prediction of binding affinity of caspase peptide inhibitors**

A binding affinity prediction for peptide inhibitors for which experimental data (Ki) have been reported [18] was performed using APF score matrix. The inhibition of caspases by peptide inhibitors is shown in Table 1. Because APF score can predict only the binding affinity of the peptide sequence in the inhibitor, the effects of Ac-, Boc- and -CHO are not taken into account. Peptide inhibitors with Ki values of $>10,000$ nM were set to the value of 10,000 nM. The correlation coefficient between log Ki values and APF score of caspase-3 was -0.69 (Figure 5a). These five peptides are not included in the analysis library (data not shown). Caspase-3 is inhibited by peptides in the following order: DEVD $>$ AEVD $>$ IETD $>$ WEHD $>$ YVAD. The APF scores of these peptides are in the same order.

The correlation coefficients between log Ki values and APF scores of caspase-7 and caspase-8 were -0.77 and -0.82, respectively (Figure 5b and 5c). When WEHD was omitted from the correlation analysis of caspase-9, the correlation coefficient was -0.86 (Figure 5d). In comparison, the APF score of LEHD, which is utilized as a selective inhibitor of caspase-9, was 0.44. This score means that LEHD has a

---

**Figure 5**

The scatter of the APF score vs log Ki for the peptides in Table 1. The solid line is the result of a linear regression fit, and $r$ denotes correlation coefficient. (a) caspase-3, (b) caspase-7, (c) caspase-8, and (d) caspase-9.
high binding affinity for caspase-9. Meanwhile, the APF score of WEHD was 0.59; therefore, it is predicted that WEHD also has a high binding affinity for caspase-9. However, the reported Ki for WEHD is not low (508 nM). APF score of a peptide is the sum of the position-specific scores of the component amino acids, and it is also assumed that each amino acid within a peptide contributes to recognition almost independently. The reason why WEHD does not have low Ki value may be improper interaction of amino acid residues in the peptide.

Construction of enriched library

The locations of inhibitory peptides (Table 1) in the sorted comprehensive peptide libraries are shown in Figure 6. All inhibitory peptides with Ki values less than 2 nM were located within the top 400 peptides in the libraries, equivalent to 5% of the comprehensive libraries. Therefore, potent peptide inhibitors are considerably enriched. Here, the enriched library is defined as the top 5% of the library. Figure 7 shows schematic overview of the APF method for constructing of enriched library.

Since the sequences of DEVD, DEPD, DELD, and DGPD, which are cleavage sites for caspase-3, also satisfy the condition of being in the top 5% of the library, it is expected that these sequences have a high binding affinity for caspase-3 (DEVD is already known as an inhibitory peptide, Ac-DEVD-CHO etc.). Meanwhile, the sequences of WEHD and YVAD have high Ki values, 1960 nM and >10,000 nM, respectively. These sequences are located at position 6506 (WEHD) and 7322 (YVAD) in the library. The sequences of DEVD, AEVD, and IETD have low Ki values for caspase-8, 0.92 nM, 1.6 nM, and 1.05 nM, respectively. The library of caspase-8 also reflects the high binding affinities of these sequences, positions of 125, 228, and 328, respectively.

We, therefore, believe that peptides with high affinities for each caspase can be efficiently screened from the enriched library as compared with comprehensive peptide libraries.

Design of a novel caspase-3 inhibitory peptide

To prove the above assumption, we tried to design a novel caspase-3 inhibitory peptide from its enriched library derived from the APF score. Initially, 400 peptides in the enriched library were docked into the active site of caspase-3 and their \( \Delta G_{\text{bind}} \) were calculated by AutoDock3.0. During the docking, the number of energy evaluations was 0.7 million, main chain angles were fixed, and other parameters were default values. Figure 8 shows the scatter of APF score vs \( \Delta G_{\text{bind}} \) for the 400 peptides from the enriched library and randomly selected 400 peptides from the non-enriched library (except the 400 peptides from the comprehensive peptide library). As expected, the 400 peptides in the enriched library had low \( \Delta G_{\text{bind}} \) (average \( \Delta G_{\text{bind}} = -15.47 \text{ kcal/mol} \)) as compared with that of peptides in the non-enriched library (average \( \Delta G_{\text{bind}} = -13.66 \text{ kcal/mol} \)).

Finally, the top 80 peptides (1% of comprehensive peptide library) with low \( \Delta G_{\text{bind}} \) in the 400 peptides of enriched library were docked into the active site of caspase-3. During the docking, the number of energy evaluations was 3 million, all of rotational torsion angles were flexible, and other parameters were default values. Ten predicted potent inhibitory peptides of caspase-3 are summarized in Table 3. The most potent peptide sequence was predicted as DEVD (\( \Delta G_{\text{bind}} = -16.15 \text{ kcal/mol} \)). DELD is the known cleavage site of D4-GDP dissociation inhibitor [25]. Seven of the ten peptides have Glu in the P3 position. Many of the well-known peptide inhibitors of caspases (see Table 1) also have Glu in the P3 position. Interestingly, only DNLD has Asn in the P3 position. Furthermore, at present, no peptide inhibitors or cleavage site in the natural substrates of caspase-3 are known to have Asn in the P3 position.

Inhibitory effect of DNLD on caspase-3 and its binding mode

To examine the inhibitory effect of DNLD on caspase-3 activity, recombinant active caspase-3 was treated with various concentrations of Ac-DNLD-CHO. As shown in Figure 9, Ac-DNLD-CHO inhibited caspase-3 in a dose-dependent manner. At 30 nM Ac-DNLD-CHO caused the inhibition of caspase-3 activity in more than 90% of the control. The IC50 value was calculated to be 3.17 nM. The potency of Ac-DNLD-CHO was slightly weak than that of a well-known peptide inhibitor, Ac-DEVD-CHO (IC50 = 1.34 nM, see Figure 9).

To obtain a better understanding of how these two peptide inhibitors interact with caspase-3, the structural model for the complex of DNLD with caspase-3 active pocket was predicted by AutoDock3.0 (Figure 10). We inspect the important interactions between the caspase-3 active site and the peptide inhibitors: (i) the Leu (P2) in DNLD forms tight hydrophobic interactions with Tyr204, Trp206, and Phe256 as well as the Val (P2) in DEVD; (ii) the Asn (P3) in DNLD forms one hydrogen bond with Ser209, but the Glu (P3) in DEVD forms two hydrogen bond with Arg207 and Ser209. These differences may affect their inhibitory potencies and/or specificities against caspase-3.

Conclusions

In the present study, we have developed a new in silico strategy based on APF score for rapid evaluation and screening of binding peptide to a target protein molecule. In the Structure-Based Virtual Screening, the computational calculation time of \( \Delta G_{\text{bind}} \) is an important
The locations of peptides (from Table 1) and sequences of cleavage sites in the sorted comprehensive peptide libraries (a) caspase-3, (b) caspase-7, (c) caspase-8, and (d) caspase-9. Hatched squares show the Enriched libraries. Columns of inhibitors show ranking, sequence, APF score and Ki, column of cleavage sites shows ranking, sequence, APF score and substrate proteins, and column of optimal sequence shows ranking, sequence, and APF score. These cleavage sites of caspase-3 have been published elsewhere [17,25,29-36]. The optimal recognition sequences are identified from PS-SCL [14].
factor as well as the computational accuracy. On a common day workstation, it is difficult to predict an appropriate binding mode and binding free energy for a peptide, which has 20 to 30 rotational bonds, against a target protein within a few minutes. Meanwhile, methods for constructing scoring functions from experimental cleavage data are rapid enough to predict binding affinities. However, many experimental data points are required to construct a precise scoring function.
The APF score developed here allows rapid prediction of peptide binding affinities. The statistical basis of APF score is derived from the calculated $\Delta G_{\text{bind}}$ between randomly extracted peptides and a target caspase using AutoDock3.0. Therefore, our system does not require experimental data for constructing a scoring function. It should be noted that the binding affinities of caspase inhibitory peptides derived rapidly from the APF score correlate well with their reported inhibitory kinetics, and that all peptides with $K_i$ values less than 2 nM were condensed in the top 5% of the sorted comprehensive peptide libraries.

To prove the utility of this method, an attractive candidate for a new caspase-3 inhibitor in the enriched library was selected by the APF score. Ac-DNLD-CHO, whose sequence is so far not known in any caspase substrate, and has a low $\Delta G_{\text{bind}}$, was chosen and its inhibitory activity
The system needs further validation; however, it is likely that by using the APF method, peptides with high affinities can be efficiently selected from the enriched library as compared to random screening. At present, small peptides have been suggested to have clinical potential, because in specific protein-protein interactions occurred in many signalling pathways, small critical domains on the molecules are revealed to be involved in their biological activities. Furthermore, many peptide ligands exist in living organisms. Thus, new computational methods for the identification of active peptides will be powerful tools for the development of new pharmaceuticals. The APF method developed here is an attractive strategy that could discover novel therapeutic agents.

Methods

Materials

Ac-DNLD-CHO was synthesized by Peptide Institute, Inc. Ac-DEVD-CHO was purchased from Peptide Institute, Inc. Ac-DEVD-pNA and recombinant human caspase-3 were from CALBIOCHEM.

Caspase-3 Inhibition assay

To determine inhibitory potency of Ac-DNLD-CHO against caspase-3, an *in vitro* caspase activity assay was performed. The activity of caspase-3 was measured using Ac-DEVD-pNA as a substrate. One unit of active recombinant human caspase-3 was incubated with peptide inhibitor at indicated concentrations in assay buffer (100 mM NaCl, 50 mM HEPES, 10 mM DTT, 1 mM EDTA, 10% glycerol, and 0.1% CHAPS, pH 7.4) at 25°C for 30 min with shaking in a 96-well assay plate (CORNING). Then, 100 µM of Ac-DEVD-pNA was added to each well and absorbance at 405 nm was measured at 25°C using a 96-well plate reader Wallace 1420 ARVOsx (PerkinElmer) at 1 min intervals for 60 min.

Preparation of caspase/inhibitory peptide complex structure

In this study, we evaluated the binding mode and affinity of caspase inhibitory peptides by AutoDock3.0, and then constructed the APF score matrices of caspases-3, -7, -8, and -9. The coordinates of caspases-3, -7, -8, and -9 were obtained from the Protein Data Bank (codes 1PAU, 1FIJ, 1QTN, and 1XQ) [26]. Since the caspase-9 inhibitor in the crystallographic structure is not a 4-amino acid peptide, LEHD, known to be an optimal sequence [14], was

| Table 3: Ten Predicted potent peptide inhibitors of caspase-3 |
|--------------------------|--------------------------|--------------------------|
| Peptide | $\Delta G_{\text{bind}}$ (kcal/mol) | APF score |
| 1 | DEVD | -16.15 | 0.73 |
| 2 | QEPD | -15.91 | 0.60 |
| 3 | TEPD | -15.75 | 0.53 |
| 4 | DELD | -15.65 | 0.59 |
| 5 | QELD | -15.63 | 0.56 |
| 6 | LEPD | -15.57 | 0.49 |
| 7 | DNLD | -15.45 | 0.52 |
| 8 | DQPD | -15.39 | 0.54 |
| 9 | DPFD | -15.16 | 0.51 |
| 10 | NEPD | -14.95 | 0.50 |
The complex structure of active site of caspase-3 with DNLD or DEVD. Nitrogen, oxygen, and carbon atoms in the active site of caspase-3 are illustrated with blue, red, and grey, respectively. (a) Predicted binding mode of DNLD is shown as yellow. (b) DEVD in crystal structure (PDB, 1PAU) is shown as orange. During the docking of DNLD against caspase-3, the number of energy evaluations was 3 million, all of rotational torsion angles were flexible, and other parameters were default values.
docked against caspase-9 using AutoDock3.0 [19], and the predicted complex structure was used for subsequent analysis. The crystallographic structures including water were deleted, and energy minimization using the TINKER software package [24] was used and energy minimization using the Amber94 forcefield was performed. During the minimization, the RMSD gradient limits were 0.01.

Construction of APF Score matrix
The feature of our computational screening strategy is construction of the Amino acid Positional Fitness (APF) score matrix for a target protein. This score matrix allow us to predict binding affinities between a target protein and peptides rapidly.

Our system comprises four steps as follows (Figure 7).

Step 1. Construction of peptide library
To identify the P2-P4 preferences of each caspase, we constructed an analysis library with aspartic acid fixed at P1 and randomized amino acids at P2-P4. In order to analyze the P1-P3-P2-D motif comprehensively, it is necessary to analyze 8000 combinations. In this system, 400 peptides equivalent to 5% of all combinations were extracted at random, and an analysis library consisting of 360 peptides and a prediction library consisting of 40 peptides were prepared. To generate initial conformation for peptides in the analysis and prediction libraries, the main chain coordinates were used for the inhibitory peptides in the crystallographic structures, and the side chains were added. Moreover, in order to exclude VDW contacts of side chains, the TINKER software package [24] was used and energy minimization using the Amber94 forcefield was performed. During the minimization, the main chain atoms were fixed and RMSD gradient limits were 0.01. Kollman atom charges were assigned and the rotational bonds of the peptides were defined with AutoTors [19].

Step 2. Calculation of binding free energy using flexible ligand docking
The \( \Delta G_{\text{bind}} \) of 400 randomly extracted peptides with a caspase were calculated.

In this study, the main chains of the peptides were fixed and the side chains were made flexible. During the docking, the number of energy evaluations was 0.7 million and other parameters were default values. The predicted binding mode of each peptide with caspase was inspected to determine whether the binding mode was appropriate. In order to arrange a functional group such as CHO on the caspase active site, a suitable binding mode of peptide is important. In generally, good docking solutions have an RMSD of 2.0 Å or less [27], and 3 peptides with Ki value of >10,000 nM do not have an RMSD of 2.0 Å or less in Table 2. Therefore, peptides with 2Å or more RMSD compared with the peptide inhibitor in the crystallographic structure were defined as not binding to caspases.

Step 3. Construction of APF Score matrix
APF score matrix based on the frequency of appearance of 20 amino acids at each of positions P1, P3, and P2 were generated in the analysis library. The APF score for amino acid i at position j is calculated as follows:

\[
APF_{ij} = \log \left( \frac{\rho_{ij}^{\text{binder}}}{\rho_{ij}^{\text{all}}} \right)
\]

where \( \rho_{ij}^{\text{binder}} \) is the frequency of amino acid i at position j among peptides in the analysis library, and \( \rho_{ij}^{\text{all}} \) is the frequency of amino acid i at position j among peptides below the threshold value (\( \Delta G_{\text{th}} \)) of binding free energy and RMSD 2 Å in the analysis library.

\[
\rho_{ij}^{\text{binder}} = \frac{c_{ij}^{\text{binder}}}{n_{\text{binder}}}
\]

\[
\rho_{ij}^{\text{all}} = \frac{c_{ij}^{\text{all}}}{n_{\text{all}}}
\]

\( c_{ij} \) denotes the number of times amino acid i appears at position j among peptides in the analysis library. \( n_{\text{all}} \) is the number of peptides in the analysis library. \( c_{ij}^{\text{binder}} \) denotes the number of times amino acid i appears at position j among peptides below the threshold value (\( \Delta G_{\text{th}} \)) of binding free energy and RMSD 2 Å in an analysis library, and \( n_{\text{binder}} \) is the number of peptides below the threshold value (\( \Delta G_{\text{th}} \)) and RMSD 2 Å. The range of position j is from 1 to 3, and corresponding to P2 to P4. The range of amino acid i is from 1 to 20, and corresponding to individual amino acids.

By using APF score matrix, it becomes possible to calculate the APF score, which shows the binding affinity of a peptide and a caspase. The APF score of a peptide is given by Equation [4].

\[
APF \text{ score } = \sum_{i=1}^{20} \sum_{j=1}^{3} c_{ij}^{\text{binder}} \cdot APF_{ij}
\]

A peptide with the P4-P3-P2-D motif is represented by a 20 × 3 matrix \( c_{ij} \) of 0s and 1s, where \( c_{ij} = 1 \) if amino acid i is at position j [28].
Correlation analysis of APF score and ΔC_{bind} of peptides consisting in the analysis library was carried out. If a high correlation is observed between APF score and ΔC_{bind}, it becomes possible to predict ΔC_{bind} using APF score. A threshold value (ΔC_{th}) was set up the value that makes the correlation coefficient between APF score and ΔC_{bind} maximal.

Next, correlation analysis between APF score and ΔC_{bind} of peptides involved in a prediction library was carried out. Here, the APF score was calculated by using APF score matrix generated from the analysis library, and the prediction capability was validated.

**Step 4. Construction of an enriched library by APF score**

It is assumed that potent peptide inhibitors are distributed randomly in the comprehensive peptide library. The APF scores of all peptides in the comprehensive peptide library were calculated using APF score matrix generated from the analysis library. Since a high APF score reflects high binding affinity, peptides in the comprehensive peptide library were sorted in descending order according to their APF score. This operation takes the high affinity peptides, which are distributed throughout the library, condenses them at a higher rank in the library. The enriched library was defined as the top 5% of the sorted comprehensive peptide library.

**Authors’ contributions**

AY implemented most of the algorithms, participated in study design, and wrote the manuscript. RT assisted with study design and performed caspase inhibition assays. ST conceived the study and participated in its design and coordination.

**Acknowledgements**

We are grateful to Dr. J. W. Ponder for providing the Tinker program and Dr. A. J. Olson for providing the AutoDock3.0 program.

**References**

1. Nicholson DW: **ICE/CED3-like proteases as therapeutic targets for the control of inappropriate apoptosis.** Nat Biotechnol 1996, 14:297-301.
2. Salvesen GS, Dixit VM: **Caspases: intracellular signaling by proteolysis.** Cell 1997, 91:443-446.
3. Cohen GM: **Caspases: the executioners of apoptosis.** Biochem J 1997, 326:1-16.
4. Nicholson DW, Thombery NA: **Caspases: killers proteases.** Nature 1998, 387:1312-1316.
5. Nicholson DW, Thombery NA: **Caspases: killer proteases.** Trends Biochem Sci 1997, 22:299-306.
6. Wellington CL, Ellerby LM, Hackam AS, Margolis RL, Triffith MA, Singara R, McCutcheon K, Salvesen GS, Propp SS, Bromm M, Rowland KJ, Zhang T, Rasper D, Roy S, Thombery N, Pinsky L, Kakiukawa A, Ross CA, Nicholson DW, Breeden DE, Hayden MR: **Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract.** J Biol Chem 1998, 273:9158-9167.
7. Kim TW, Pettingell WH, Jung YK, Kovacs DM, Tanzi RE: **Alternative cleavage of Alzheimer-associated presenilins during apoptosis by a caspase-3 family protease.** Science 1997, 277:373-376.
8. Gervais FG, Xu D, Robertson GS, Vaillancourt JP, Zhu Y, Huang J, LeBlanc A, Smith D, Rigby M, Shearman MS, Clarke EE, Zheng H, Van Der Pliegh L, Ruffolo SC,(article) Thornberry NA, Xanthisoulos S, Zamboni RJ, Roy S, Nicholson DW: **Involvement of caspases in proteolytic cleavage of Alzheimer’s amyloid-beta precursor protein and amyloidogenic A beta peptide formation.** Cell 1999, 97:395-406.
9. Thompson CB: **Apoptosis in the pathogenesis and treatment of disease.** Science 1995, 267:1456-1462.
10. Wilson KP, Black JF, Thomson JA, Kim EE, Griffith JP, Nava MA, Murcko MA, Chambers SP, Aldape RA, Raybuck SA, Livingston DJ: **Structure and mechanism of interleukin-1 beta converting enzyme.** Nature 1994, 368:270-275.
11. Walter NPC, Talanian RV, Brady KD, Dang LC, Fazil KM, Gallant M, Labelle M, Peterson EP, Rasper DM, Ruel R, Vaillancourt JP, Thombery NA, Benarow JR: **The three-dimensional structure of apopain:** CPP32, a key mediator of apoptosis. Nature Struct Biol 1996, 3:619-625.
12. Mittl PRE, Di Marco S, Krebs JF, Bai X, Karanewske DS, Priestle JP, Tomasselli KJ, Grutter MG: **Structure of recombinant human CPP32 in complex with the tetrapeptide acetyl-Asp-Val-Ala-Asp fluoromethyl ketone.** J Biol Chem 1997, 272:6539-6547.
13. Thombery NA, Rano TA, Peterson EP, Rasper DM, Timkey T, Garcia-Calvo M, Houzager VM, Nordstrom PA, Roy S, Vaillancourt JP, Chapman KT, Nicholson DW: **A combinatorial approach defines specificities of members of the caspase family and granzyme B.** Functional relationships established for key mediators of apoptosis. J Biol Chem 1997, 272:17907-17911.
14. Humke EW, Ni J, Dixit VM: **ERICE, a novel FLICE-activatable caspase.** J Biol Chem 1998, 273:15702-15707.
15. Shimim H, Snipas SJ, Vincenz C, Salvesen G, Dixit VM: **Caspase-14 is a novel developmentally regulated protease.** J Biol Chem 1998, 273:29648-29653.
16. Lazebnik YA, Kaufmann SH, Desnoyers S, Poirier GG, Earnshaw WC: **Cleavage of poly(ADP-ribose) polymerase by a proteinase with properties like ICE.** Nature 1994, 371:334-347.
17. Garcia-Calvo M, Houzager VM, Nordstrom PA, Roy S, Vaillancourt JP, Chapman KT, Nicholson DW: **A combinatorial approach defines specificities of members of the caspase family and granzyme B.** Functional relationships established for key mediators of apoptosis. J Biol Chem 1997, 272:17907-17911.
18. Chou KC, Tomasselli AG, Reardon IM, Heinrikson RL: **Predicting human immunodeficiency virus protease cleavage sites in proteins by a discriminant function method.** Proteins 1996, 24:51-72.
19. Poorman RA, Tomasselli AG, Heinrikson RL, Kazdy FJ: **A cumulative specificity model for proteases from human immunodeficiency virus types 1 and 2, inferred from statistical analysis of an extended substrate data base.** J Biol Chem 1991, 266:14554-14561.
20. Rosin CD, Belew RK, Morris GM, Olson AJ, Goodsell DS: **Coevolutionary analysis of resistance-evading peptidomimetic inhibitors of HIV-1 protease.** Proc Natl Acad Sci U S A 1999, 96:1369-1374.
21. Hetenyi C, van der Spoel D: **Efficient docking of peptides to proteins without prior knowledge of the binding site.** Protein Sci 2002, 11:7729-7737.
22. Ponder JW, Richards FM: **An Efficient Newton-like Method for Molecular Mechanics Energy Minimization of Large Molecules.** J Comput Chem 1987, 8:1016-1024.
23. Goldberg YP, Nicholson DW, Rasper DM, Kalchman MA, Koidhe HB, Graham RK, Bromm M, Kazemi-Esfarjani P, Thombery NA, Vaillancourt JP, Hayden MR: **Cleavage of huntingtin by apopain, a proapoptotic cysteine protease, is modulated by the polyglutamine tract.** Nat Genet 1996, 13:442-449.
24. The RCSB Protein Data Bank: [http://www.rcsb.org/pdb/](http://www.rcsb.org/pdb/)
27. Nissink JW, Murray C, Hartshorn M, Verdonk ML, Cole JC, Taylor R: A new test set for validating predictions of protein-ligand interaction. Proteins 2002, 49:457-71.

28. Zhao Y, Gran B, Pinilla C, Markovic-Plese S, Hemmer B, Tzou A, Whitney LW, Biddison WE, Martin R, Simon R: Combinatorial peptide libraries and biometric score matrices permit the quantitative analysis of specific and degenerate interactions between clonotypic TCR and MHC peptide ligands. J Immunol 2001, 167:2130-41.

29. Sakahira H, Enari M, Nakata S: Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis. Nature 1998, 391:96-99.

30. Song Q, Lees-Miller SP, Kumar S, Zhang Z, Chan DW, Smith GC, Jackson SP, Alnemri ES, Litwack G, Khanna KK, Lavin MF: DNA-dependent protein kinase catalytic subunit: a target for an ICE-like protease in apoptosis. EMBO J 1996, 15:3238-3246.

31. Waterhouse N, Kumar S, Song Q, Strike P, Sparrow L, Dreyfuss G, Alnemri ES, Litwack G, Lavin M, Watters D: Heteronuclear ribonucleoproteins C1 and C2, components of the spliceosome, are specific targets of interleukin 1 beta-converting enzyme-like proteases in apoptosis. J Biol Chem 1996, 271:29335-29341.

32. Emoto Y, Manome Y, Meinhardt G, Kishki H, Kharbanda S, Robertson M, Ghayur T, Wong WW, Kamen R, Lavin MF: DNA-dependent protein kinase catalytic subunit: a target for an ICE-like protease in apoptotic cells. J Exp Med 1995, 184:2599-2604.

33. Wang ZQ, Auer B, Stingl L, Berghammer H, Haider D, Schweiger M, Wagner EF: Mice lacking ADPRT and poly(ADP-ribose)lation develop normally but are susceptible to skin disease. Genes Dev 1995, 9:509-520.

34. Browne SJ, Williams AC, Hague A, But A, Paraskeva C: Loss of APC protein expressed by human colonic epithelial cells and the appearance of a specific low-molecular-weight form is associated with apoptosis in vitro. Int J Cancer 1994, 59:56-64.

35. An B, Dou QP: Cleavage of retinoblastoma protein during apoptosis: an interleukin 1 beta-converting enzyme-like protease as candidate. Cancer Res 1996, 56:438-442.