Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

In silico identification and crystal structure validation of caspase-3 inhibitors without a P1 aspartic acid moiety

Rajkumar Ganesan,‡ Stjepan Jelakovic, Peer R. E. Mittl,* Amedeo Caflisch* and Markus G. Grütter

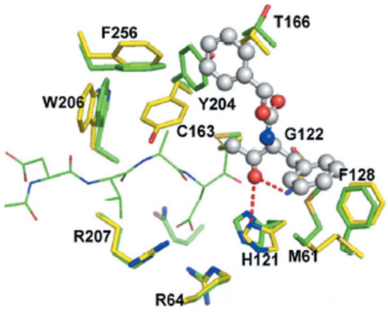
Department of Biochemistry, University of Zürich, Winterthurer Strasse 190, 8057 Zürich, Switzerland

‡ Present address: Department of Antibody Discovery and Protein Engineering, MedImmune Inc., 1 MedImmune Way, Gaithersburg, MD 20878, USA.

Correspondence e-mail: mittl@bioc.uzh.ch, caflisch@bioc.uzh.ch

Received 24 January 2011
Accepted 16 May 2011

PDB References: caspase-3–compound **1**, 2xyg; caspase-3–compound **5**, 2xyh; caspase-3–compound **8**, 2xyp.



Using a fragment-based docking procedure, several small-molecule inhibitors of caspase-3 were identified and tested and the crystal structures of three inhibitor complexes were determined. The crystal structures revealed that one inhibitor (NSC 18508) occupies only the S1 subsite, while two other inhibitors (NSC 89167 and NSC 251810) bind only to the prime part of the substrate-binding site. One of the major conformational changes observed in all three caspase-3–inhibitor complexes is a rotation of the Tyr204 side chain, which blocks the S2 subsite. In addition, the structural variability of the residues shaping the S1–S4 as well as the S1′ subsites supports an induced-fit mechanism for the binding of the inhibitors in the active site. The high-resolution crystal structures reported here provide novel insights into the architecture of the substrate-binding site, which might be useful for the design of more potent caspase inhibitors.

1. Introduction

Caspases form a unique class of cysteine aspartate-specific proteases according to their substrate specificities and biological functions (Grütter, 2000; Fuentes-Prior & Salvesen, 2004). The inhibition of caspases represents a highly promising avenue for intervention in a number of conditions involving apoptosis-mediated cell and tissue damage in acute and chronic diseases (O'Brien & Lee, 2004). Many potent caspase inhibitors have been synthesized based on the structures of caspase peptide substrates. However, because of their peptidyl structure most of these inhibitors do not possess sufficient cell permeability to be considered as candidates for drug development.

12 human caspases are known and they demonstrate a stringent specificity for an aspartic acid residue in the P1 position. They use their active-site thiol to cleave peptide/protein targets exclusively C-terminal to an aspartic acid residue. Furthermore, caspases are selective for residues in the S4 and S1′ subsites (Ganesan, Jelakovic *et al.*, 2006; Ganesan, Mittl *et al.*, 2006; Goode *et al.*, 2005). Caspase inhibitors typically contain an electrophilic group termed the 'warhead', which reacts with the nucleophilic active-site cysteine residue. The most commonly used warheads are halomethyl ketones and aldehydes. Coupled with the preferred tetrapeptide sequence, they produce potent inhibitors. Recently, two new classes of aza-peptides have been identified as specific warheads for caspases (James *et al.*, 2004; Ekici *et al.*, 2006). Active caspases are $(\alpha\beta)_2$ heterotetramers that can be described as homodimers of heterodimeric subunits with two active sites located in close proximity to the dimer interface (Wilson *et al.*, 1994; Rotonda *et al.*, 1996; Mittl *et al.*, 1997; MacCorkle *et al.*, 1998; Watt *et al.*, 1999; Schweizer *et al.*, 2003; Hardy *et al.*, 2004; Fuentes-Prior & Salvesen, 2004; McCluggage *et al.*, 2006; Ganesan, Mittl *et al.*, 2006). Thermodynamic analysis and NMR experiments revealed that residues from the dimer interface are critical for stabilizing the active-site loops (Piana *et al.*, 2003; Keller *et al.*, 2009, 2010). To identify novel scaffolds for the design of caspase-directed inhibitors, the active-site pocket of caspase-3 was probed using a fragment-based docking procedure and identified compounds were analyzed *in vitro* using biochemical and structural biology techniques.

2. Materials and methods

2.1. Computational screening

We used *in silico* screening to identify compounds containing halomethyl ketone or aldehyde groups that can form a covalent bond to the catalytic cysteine residue of caspases. The computational approach used presumes three fragments for the placement of a flexible ligand in the binding site. In fragment-based docking approaches, rigid or almost rigid molecular fragments are automatically placed in the binding site and used as anchors for guiding the docking of the compounds from which they originate (Cecchini *et al.*, 2004). In brief, the first program of the docking suite, *DAIM* (*Decomposition and Identification of Molecules*), decomposes a ligand into substructures and prioritizes the resulting fragments according to their suitability as anchor fragments for use in an initial rigid docking procedure (Kolb & Caflisch, 2006). Fragments that form highly favourable interactions with the protein upon binding constitute the most suitable anchor fragments for fragment-based docking. *DAIM* fingerprints, which include fields that consider size and functional group features, are used for the selection of the most suitable anchor fragments. Optimal positions and orientations for the small to medium-sized fragments generated by *DAIM* are determined by *SEED* (*Solvation Energy for Exhaustive Docking*; Majeux *et al.*, 1999, 2001). This program identifies optimal positions and orientations of small to medium-sized molecular fragments in the binding site of the protein. The fragments are then docked exhaustively by placing each fragment in several thousand different positions with multiple orientations and the binding energy for each fragment pose is estimated whenever severe clashes are not present. Polar fragments are placed under the condition that at least one hydrogen bond is formed to the protein. Apolar fragments are docked into hydrophobic regions of the binding site. Binding energies are calculated as the sum of van der Waals and electrostatic terms, taking into account the interaction energies between the fragment and the protein as well as desolvation energies. In the last step of the docking procedure, the whole ligand molecule is docked flexibly into the binding pocket by the program *FFLD* (*Fast Flexible Ligand Docking*; Budin *et al.*, 2001; Cecchini *et al.*, 2004). This program uses a genetic algorithm for modification of the torsion angles and applies an efficient scoring function (Scarsi *et al.*, 1999). The placement of the ligand in the binding site is defined by the anchor points and a least-squares fitting procedure, modifying only the conformation of the ligand. The position and orientation of the ligand in the binding site are determined by the best binding modes of its fragments that were initially docked by *SEED*. Since the ligand is not covalently linked to the active site during the *FFLD* run, the program does not have to account for formation of the covalent bond. For efficiency reasons, the scoring function used in *FFLD* does not explicitly include solvation and is based solely on van der Waals and hydrogen-bond energy terms. Nevertheless, solvation effects are indirectly considered by utilizing the best binding modes previously determined by *SEED*. Poses with the most favourable *FFLD* energies are further minimized in the rigid protein using the CHARMM22 force field (Accelrys Inc.) and during this step the ligand is covalently linked to the active-site cysteine of caspase-3.

In order to partially take into account protein flexibility induced by ligand binding, the crystal structures of four different caspase-3–inhibitor complexes (PDB entries 1cp3, 1nme, 1nmq and 1nms) were used as templates for the docking experiments (Choong *et al.*, 2002; Mittl *et al.*, 1997; Schweizer *et al.*, 2003; Watt *et al.*, 1999; Erlanson *et al.*, 2003). For the docking experiments the following 26 amino acids, which are located within a cutoff radius of 5 Å from any ligand atom, were selected as binding-site residues: Met61–Ser65, Ser120–Glu123, Gln161–Cys163, Thr166, Tyr204–Ser209, Trp214, Ser249–Asp253 and Phe256. The defined binding-site residues provide the anchor points for *SEED* (Cecchini *et al.*, 2004). All crystallographic water molecules were removed from the protein structures. H atoms, considering appropriate ionization states of both the acidic and basic amino-acid residues, atom types and partial charges were assigned to the protein targets and the ligands using the programs *BABEL* (Guha *et al.*, 2006), *WITNOTP* (A. Widmer, Novartis AG, Basel, Switzerland) and a method based on the partial equalization of orbital electronegativity (No, Grant, Jhon *et al.*, 1990; No, Grant & Scheraga, 1990). All histidine residues, except the active-site His121, were considered neutral. The crystal structures were then energy-minimized using the CHARMM22 force field (Accelrys Inc.). During the refinement, heavy atoms were fixed and the H atoms were kept flexible in order to find suitable hydrogen positions. Subsequently, the covalently bound ligands were removed from the protein structures and the active-site Cys163 was ionized.

Filtering the Available Chemical Directory database (240 000 compounds), the National Cancer Institute (NCI) 3D database (140 000 compounds) and the Cambridge Crystallographic Data Centre database (250 000 compounds) yielded a total of about 7200 molecules containing halomethyl ketone or aldehyde groups. To facilitate crystallization of caspase-3–inhibitor complexes, only the irreversibly binding halomethyl ketones (about 650 compounds) were used for flexible-ligand docking. Prior to screening the molecules from the three different databases, a docking protocol was established by docking four known halomethyl ketone inhibitors into the active sites of the respective caspase-3 protein structure templates (PDB entries 1cp3, 1nme, 1nmq and 1nms). In order to remove any bias originating from the ligand conformation observed in the crystal structures, the four ligands were minimized outside the active site before starting the docking procedure. During the *FFLD* run the fragment containing the electrophilic carbon was placed in the vicinity of the active-site Cys163 to allow the formation of a covalent bond to the protein. Nonetheless, the fragment was not fixed during the flexible docking procedure. Each ligand was docked into the active site of each of the four caspase-3 conformations and ten genetic algorithm runs were performed for every docking experiment. For each run, the ligand pose with the lowest *FFLD* energy was post-processed by minimization using the CHARMM22 force field (Accelrys Inc.), while keeping protein non-H atoms rigid. During this post-processing step the ligands were linked covalently to the active-site Cys163. Protein–ligand complexes were minimized by the conjugate-gradient algorithm with an energy-convergence criterion of 0.004 kJ mol^{-1} Å^{-1}. In order to prevent artificial deviations because of vacuum effects and to crudely approximate solvent effects, the electrostatic energy term was screened by a distance-dependent dielectric function ($\varepsilon = 4r$, where r is the distance between partial charges) and the default nonbonding cutoff of 14 Å was used during the refinement. The accuracies of the redocking and cross-docking experiments were assessed by calculating the root-mean-square deviations (r.m.s.d.s) for all non-H ligand atoms between the docked conformations and the conformations observed in the experimental crystal structures. Only the best-scoring orientation from each docking run was considered.

The 650 halomethyl ketone compounds were docked and post-processed using the same protocol as described for the four known ligands. Again, the protein coordinates of the four caspase structures 1cp3, 1nme, 1nms and 1nmq were used as templates for docking. The halogen atoms of the halomethyl ketone groups of the molecules retrieved from the databases were removed prior to docking. Ten

genetic algorithm runs were performed for the docking of each ligand into the active sites of the four caspase-3 structures. Compounds with plausible poses were either obtained from the Drug Synthesis and Chemistry Branch, the Developmental Therapeutics Program, NCI or purchased from commercial suppliers and tested in an *in vitro* activity assay for their inhibition potency.

2.2. Caspase purification

Human caspase-3 was produced in *Escherichia coli* as inclusion bodies, refolded and purified as described previously (Garcia-Calvo *et al.*, 1999). Both subunits of caspase-3 were expressed separately in *E. coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene). Cultures (500 ml) were grown to a density of A_{600} = 0.5 at 310 K in Luria–Bertani medium. Expression was induced by the addition of IPTG (1 m*M*) and the culture was shaken at 310 K for 4 h post-induction. Cells were harvested, washed in PBS buffer and lysed using a French press to retrieve the protein localized in the inclusion-body portion. Refolding was achieved by rapid dilution of equimolar amounts of each subunit to a final concentration of about 100 µg ml^{-1} per subunit. The protein was purified by anion-exchange chromatography (Resource-Q, GE Healthcare). Active caspase was further purified by size-exclusion chromatography (Superdex S200, GE Healthcare).

2.3. Kinetic analysis

The inhibitory activities of the compounds were measured at 10 n*M* caspase-3 concentration. All compounds used for testing were dissolved at 50 m*M* in DMSO prior to assay. Typically, the enzyme was incubated with the compound at a concentration of 100 µ*M* for 30 min at 310 K. The remaining caspase activity was measured by following the increase in fluorescence using the substrate acetyl-DEVD-7-amido-4-methylcoumarin. Caspase activity was determined from the initial rate of hydrolysis by measuring the increase in fluorescence at an excitation wavelength of 360 nm and an emission wavelength of 460 nm for 15–30 min at 310 K using an HTS 7000 Plus Bio Assay plate reader (Perkin Elmer). Both the enzyme and the substrate were diluted in assay buffer consisting of 20 m*M* piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES) pH 7.2, 100 m*M* NaCl, 10 m*M* 1,4-dithiothreitol (DTT), 1 m*M* EDTA, 0.1%(*w*/*v*) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) and 10%(*w*/*v*) sucrose. The active caspase concentration was determined by active-site titration using the covalently binding inhibitor benzyloxycarbonyl-DEVD-chloromethylketone. As a prerequisite for the assay, the concentration of DMSO was less than 2% and the substrate:enzyme ratio was at least 50:1. The assay volume was 100 µl in a 96-well microtitre plate.

2.4. Crystallization and structure determination

For structural studies, the purified caspase-3 was inhibited with a threefold molar excess of the inhibitor. Crystals of caspase-3–inhibitor complexes were grown by mixing equal volumes of protein solution (10 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0, 10 m*M* DTT) and reservoir solution [0–15%(*w*/*v*) polyethylene glycol 6000, 100 m*M* sodium citrate pH 5.0]. Crystallization conditions were similar to previously reported conditions (Ganesan, Mittl *et al.*, 2006). Prior to data collection, crystals were frozen in a nitrogen stream after a short soak in reservoir solution containing 20% glycerol. The crystals belonged to space group *I*222. X-ray data for the protein–inhibitor **1** complex were collected at the Swiss Light Source synchrotron (Paul Scherrer Institute, Villigen, Switzerland). Data processing was performed with *XDS* (Kabsch, 2010). Diffraction data for the complexes with the inhibitors **5** and **8** were collected using a rotating-anode generator (Bruker Nonius FR591) at 100 K. Images were integrated with *MOSFLM* and scaled with *SCALA* from the *CCP*4 suite of programs (Winn *et al.*, 2011). The structures were solved by the difference Fourier technique. Model building was performed with the program *O* (Jones *et al.*, 1991) and structure refinement was performed using *CNS* (Brünger *et al.*, 1998). The topology and parameter files for the energy-minimized inhibitor were generated

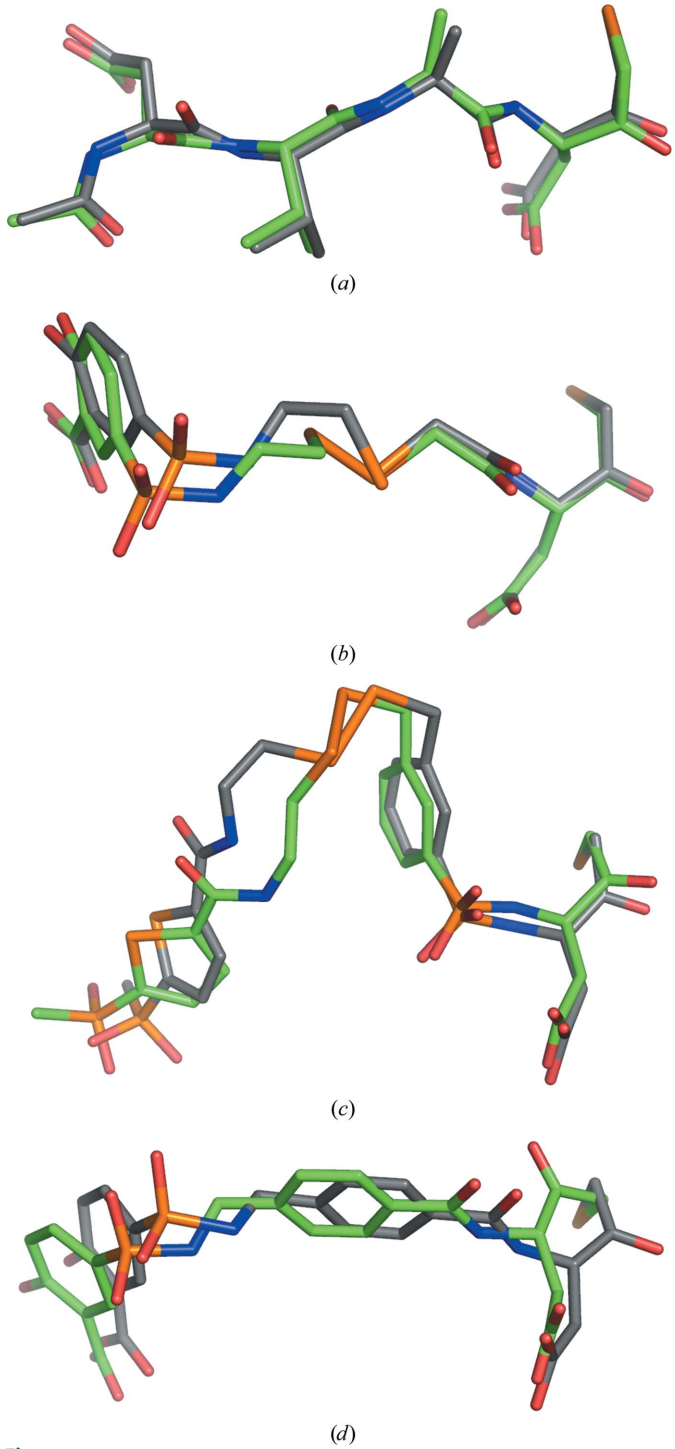


Figure 1
Comparison of the redocked poses (green C atoms) with the corresponding conformations observed in the known caspase-3–inhibitor structures (grey C atoms) from PDB entries 1cp3 (*a*), 1nme (*b*), 1nmq (*c*) and 1nms (*d*).

using *XPLO2D*. Unit-cell parameters are as listed in Table 3. The final crystallographic *R* and free *R* factors were in the ranges 17–18% and 19–22%, respectively. The final models consisted of residues 29–174 of the α-subunit and residues 186–277 of the β-subunit for all three caspase-3–inhibitor complexes described here. The majority of the molecular-graphics figures were prepared with *PyMOL* (DeLano, 2002), whereas Fig. 6 was prepared with *Maestro* (Schrödinger LLC, New York, USA). Coordinates and structure-factor amplitudes for the three caspase-3–inhibitor complexes have been deposited in the PDB under the accession codes 2xyg (caspase-3–compound **1**), 2xyh (caspase-3–compound **5**) and 2xyp (caspase-3–compound **8**).

3. Results and discussion

Here, we present the identification and structural characterization of small-molecule inhibitors binding to the substrate-binding site of caspase-3 using an *in silico* screening approach. After validation of the docking procedure by redocking and cross-docking of structurally characterized caspase-3 inhibitors using the fragment-based docking suite *DAIM–SEED–FFLD* (Kolb & Caflisch, 2006; Majeux *et al.*, 1999, 2001; Budin *et al.*, 2001), a substructure search for halomethyl ketone derivatives in three different databases was carried out.

The results of the redocking experiments, in which each of the four ligands was docked back into its cocrystallized protein conformation (PDB entries 1cp3, 1nme, 1nms and 1nmq), are shown in Fig. 1. The redocked binding modes for all four ligands were very similar to the experimentally observed binding modes. The r.m.s.d.s between the crystallographic coordinates and the docked ligands were 0.5, 1.1, 1.2 and 1.4 Å for 1cp3, 1nme, 1nmq and 1nms, respectively.

Cross-docking, in which each ligand was docked into the same protein but using a different crystal structure, was successful in six out of 12 cases when considering an r.m.s.d. value of below 2.4 Å as a successfully docked ligand conformation (Table 1; Fig. 2). The ligand superpositions used for r.m.s.d. calculations were based on the superimposed protein C^{α} atoms. While the ligand from the crystal structure 1nmq could be accurately redocked with an r.m.s.d. between the experimental and modelled ligand of 1.2 Å, cross-docking of this ligand failed in all three cases, perhaps because of the conformational changes in the protein upon ligand binding. Moreover, the 1cp3 protein structure does not seem to accommodate the three cross-docked inhibitors, which is probably a consequence of the different inhibitor scaffolds. Structure 1cp3 contains a peptidic inhibitor, whereas all other cross-docked inhibitors are clearly nonpeptidic except for the P1 Asp.

The fact that docking of some ligands into the binding site of a specific protein conformation can be difficult in some cases indicates that different experimentally determined protein conformations should be used in docking experiments, especially when the conformation of active-site residues is altered upon ligand binding. In the absence of experimentally determined structural variability, protein conformations generated by Monte Carlo (Caflisch *et al.*, 1997) or molecular-dynamics (Ekonomiuk *et al.*, 2009) simulations might be a choice for sampling different conformations of the protein target.

Visual inspection of the best scored poses of our docking results led to a subset of 27 compounds which were subsequently ordered either from the NCI Open Chemical Repository or from other commercial suppliers. Only 21 of these compounds could be obtained commercially and were subsequently assayed for inhibitory activity against caspase-3. Although all of these compounds showed some inhibitory activity, eight showed more than 50% inhibition in the caspase activity assay when tested at a concentration of 100 μM. The

Table 1
Results of redocking (diagonal from top left to bottom right) and cross-docking simulations of methyl ketone inhibitors containing an irreversible fluoromethyl ketone (i1cp3) or arylacyloxymethyl ketone warhead (i1nme, i1nmq and i1nms).

R.m.s.d. values for non-H atoms between the docked conformation of most favourable energy and the conformation from the crystal structure are given (in Å). The ligand superpositions used for r.m.s.d. calculations are based on the superimposed protein C^{α} atoms. Ligand conformations with an r.m.s.d. value below 2.4 Å are considered as successfully docked poses. The structures of the inhibitors are shown in Fig. 2.

	1cp3	1nme	1nmq	1nms
i1cp3	0.5	1.0	2.2	4.0
i1nme	3.8	1.1	1.3	1.5
i1nmq	4.4	5.7	1.2	3.6
i1nms	3.9	1.4	1.1	1.4

Table 2
Active-site-directed caspase-3 inhibitors.

The activity refers to the percentage activity remaining after incubation, as described in §2.3. The structures of the inhibitors are shown in Fig. 3.

Compound	NSC/CAS Nos.	Activity (%)
1	NSC 89167, CAS 329-30-6	71
2	NSC 24873, CAS 1775-46-8	69
3	NSC 107152, CAS 10161-88-3	68
4	NSC 211621, CAS 19159-87-6	66
5	NSC 18508, CAS 60254-71-9	59
6	NSC 106024, CAS 10161-89-4	53
7	NSC 107443, CAS 37660-62-1	53
8	NSC 251810, CAS 26049-94-5	50

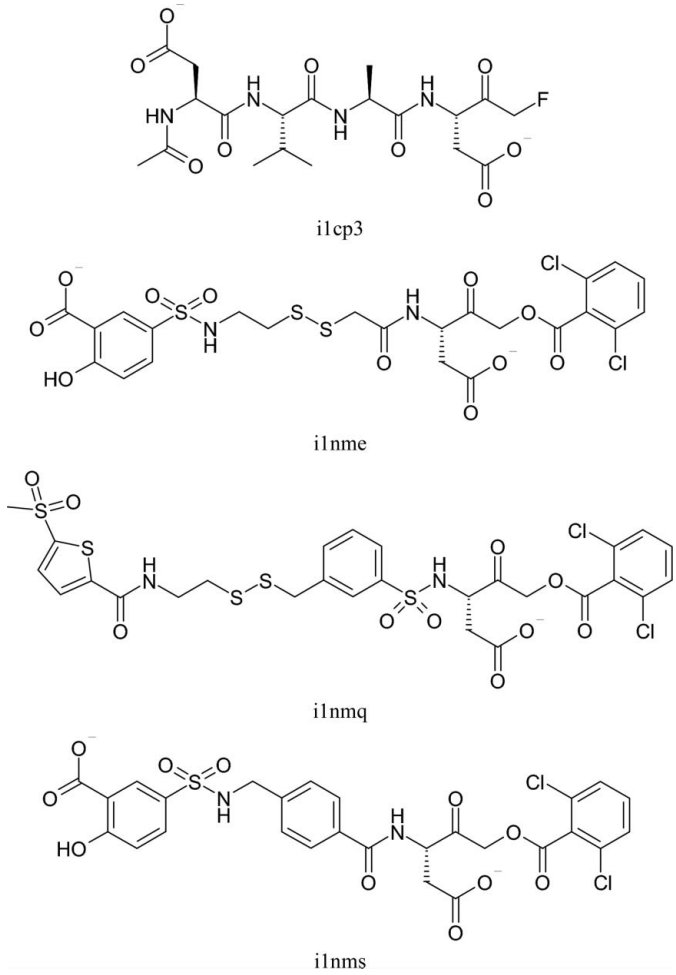


Figure 2
Structures of i1cp3, i1nme, i1nmq and i1nms.

relative inhibition data for these compounds and their chemical structures are given in Table 2 and Fig. 3, respectively.

All 21 identified inhibitors were screened for cocrystallization with caspase-3. However, only compounds **1**, **5** and **8** yielded crystals suitable for X-ray structure determination. A summary of the crystallographic data-processing and refinement statistics is given in Table 3. The overall conformations of caspase-3 reported here are very similar to the conformations of previously determined caspase-3–inhibitor complexes (Mittl *et al.*, 1997; Rotonda *et al.*, 1996). The halomethyl ketone warhead groups are covalently bound to the thiol group of the active-site cysteine through a thioether bond. Compound **5** mimics an aspartic acid coupled to a methylene chloride moiety at the C-terminus. The binding mode is very similar to that of inhibitors possessing a P1 aspartate moiety (Mittl *et al.*, 1997; Rotonda *et al.*, 1996). This ligand occupies the S1 subsite and is stabilized by ionic interactions with Arg64 and Arg207 as well as a hydrogen-bond interaction with Gln161 (Fig. 4*a*). In the known caspase-3–inhibitor structures there are two critical intermolecular main chain–main chain interactions. One hydrogen bond is observed between the P1 amide of the inhibitor and the backbone carbonyl of the conserved Ser205, while a second hydrogen bond is formed between the P3 carbonyl group and the main-chain amide group of the conserved Arg207. The lack of the amide group in compound **5** and thus the missing hydrogen-bond interaction renders the methylene C atoms of compound **5** very flexible, as indicated by high *B* values. The carbonyl group of the inhibitor is stabilized by the 'oxyanion hole' formed by the amide groups of Cys163 and Gly122. The inhibitor occupies the S1 subsite, thus leaving subsites S2–S4 unoccupied. As a consequence, the residues involved in shaping these

Table 3
X-ray data-collection and refinement statistics of caspase-3–inhibitor crystal structures.

Values in parentheses are for the highest resolution shell. All data were collected at 100 K.

	Compound **1**, NSC 89167	Compound **8**, NSC 251810	Compound **5**, NSC 18508
Space group	*I*222		
Unit-cell parameters			
a (Å)	68.7	69.5	68.5
b (Å)	83.5	83.7	83.5
c (Å)	96.1	95.8	95.8
Resolution range (Å)	20–1.57 (1.65–1.57)	20–1.86 (1.95–1.86)	20–1.89 (2.00–1.89)
No. of unique reflections	37129 (2840)	23519 (3379)	21032 (3029)
$\langle I/\sigma(I)\rangle$	30.5 (4.3)	27.8 (13.8)	21.1 (8.2)
R_{merge} (%)	9.8 (22.0)	4.2 (10.6)	5.0 (16.9)
Completeness (%)	90.3 (88.4)	99.3 (98.5)	95.6 (94.5)
Multiplicity	7.1 (5.2)	5.1 (5.1)	4.4 (4.4)
Refinement			
Resolution range (Å)	10–1.58 (1.60–1.58)	10–1.86 (1.90–1.86)	10–1.89 (1.93–1.89)
R factor (%)	17.9 (25.7)	16.7 (21.0)	17.7 (20.7)
R_{free} (%)	19.7 (27.6)	20.4 (27.9)	21.6 (22.5)
No. of reflections			
Used for refinement	35638	22609	20081
Used for calculation of R_{free}	1491	920	820
R.m.s.d. bonds (Å)	0.007	0.007	0.007
R.m.s.d. angles (°)	1.445	1.411	1.411
Average *B* (Å^2)			
Protein	11.5	11.5	13.3
Ligand atoms	34.6	44.5	32.8
Water molecules	30.5	31.1	28.9
No. of atoms			
Protein	1929	1929	1929
Ligand atoms	22	22	8
Water molecules	426	400	297

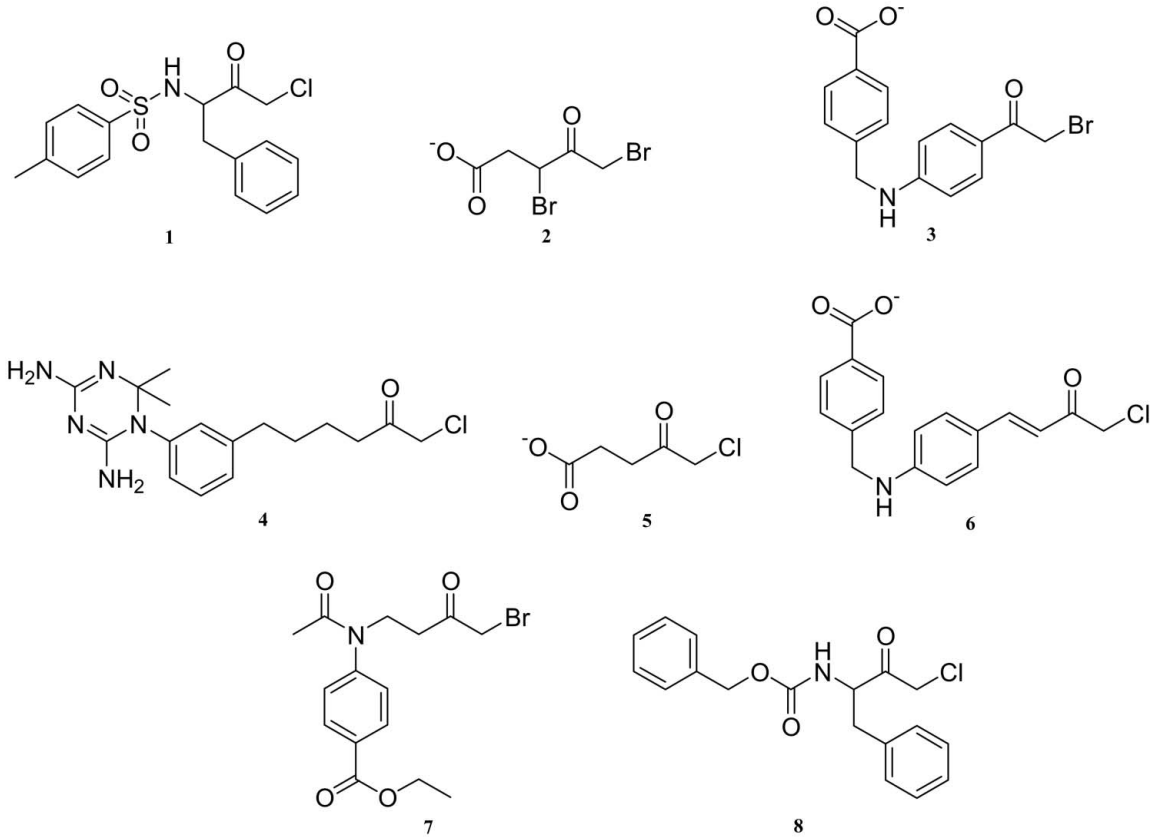


Figure 3
Structures of active-site-directed caspase-3 inhibitors. (**1**) *N*-(4-chloro-3-oxo-1-phenyl-2-butyl)-4-methylbenzenesulfonamide; (**2**) 3,5-dibromo-4-oxopentanoate; (**3**) 4-{[4-(2-bromoacetyl)anilino]methyl}benzoate; (**4**) 1-chloro-6-[3-(4,6-diamino-2,2-dimethyl-1,2-dihydro-1,3,5-triazin-1-yl)phenyl]hexan-2-one; (**5**) 5-chloro-4-oxopentanoate; (**6**) 4-{[4-(4-chloro-3-oxobut-1-enyl)anilino]methyl}benzoate; (**7**) ethyl 4-[*N*-(4-bromo-3-oxobutyl)acetamido]benzoate; (**8**) benzyl *N*-(4-chloro-3-oxo-1-phenyl-2-butyl)carbamate.

sites have different orientations compared with those observed in other known structures (Mittl *et al.*, 1997; Rotonda *et al.*, 1996). One significant change is observed for Tyr204, which rotates approximately 90° around χ_1 and occupies the S2 pocket, where it is stabilized by hydrophobic contacts with Cys163, Trp206 and Phe256 (Fig. 4*a*).

In contrast to the mode of binding observed in the caspase-3–compound **5** complex, compound **1** binds only in the prime site of caspase-3. This unusual binding in the prime site renders the substrate-binding subsites S1–S4 unoccupied. Apart from some movements of the side-chain conformation of the residues forming these subsites, a similar conformational change of Tyr204 as observed in the case of the caspase-3–compound **5** complex was also observed in this structure. Given the lack of the critical P1 Asp residue, the S1 subsite is occupied by a cluster of water molecules in the caspase-3–compound **1** structure. As a consequence, the active-site Arg207 is displaced by about 0.5 Å in comparison to the caspase-3–compound **5** complex and it adopts a slightly different side-chain orientation. Compound **1** is covalently bound to the active-site cysteine residue and its P1 carbonyl group is stabilized partly by the 'oxyanion hole' (the amide group of Gly122; Figs. 4*b* and 5*a*). The phenylalanine side chain of the inhibitor covers the S1′ subsite and is involved in hydrophobic contacts with Met61 and Phe128, whereas the *p*-toluoyl group is stabilized by the hydrophobic parts of the Thr166 and Glu123 side chains. As usually observed in arylsulfonamide structures, the conformation of the N—C bond in the aryl-SO_2-NH-C segment adopts a *gauche* torsion with respect to one of the S═O bonds. The sulfone amide group forms water-mediated interactions with the main-chain carbonyl groups of Gly122 and Gly165. Therefore, the sulfone amide could presumably be replaced by other groups linking the *p*-toluoyl and benzyl groups of the inhibitor.

The binding of the carbamic acid group in the structure of the caspase-3–compound **8** complex is reminiscent of the binding of the sulfone amide group of compound **1**. However, the *p*-toluoyl and benzyloxy groups of compounds **1** and **8**, respectively, bind to completely different regions of the active site. The benzyloxy group of compound **8** is well placed in a deep hydrophobic cavity delineated by Thr166, Leu168, Tyr204 and Phe256, while the phenylalanine side chain is stabilized by hydrophobic contacts with Met61 and Phe128 (Figs. 4*c* and 5*b*). The carbonyl group of the inhibitor is stabilized by

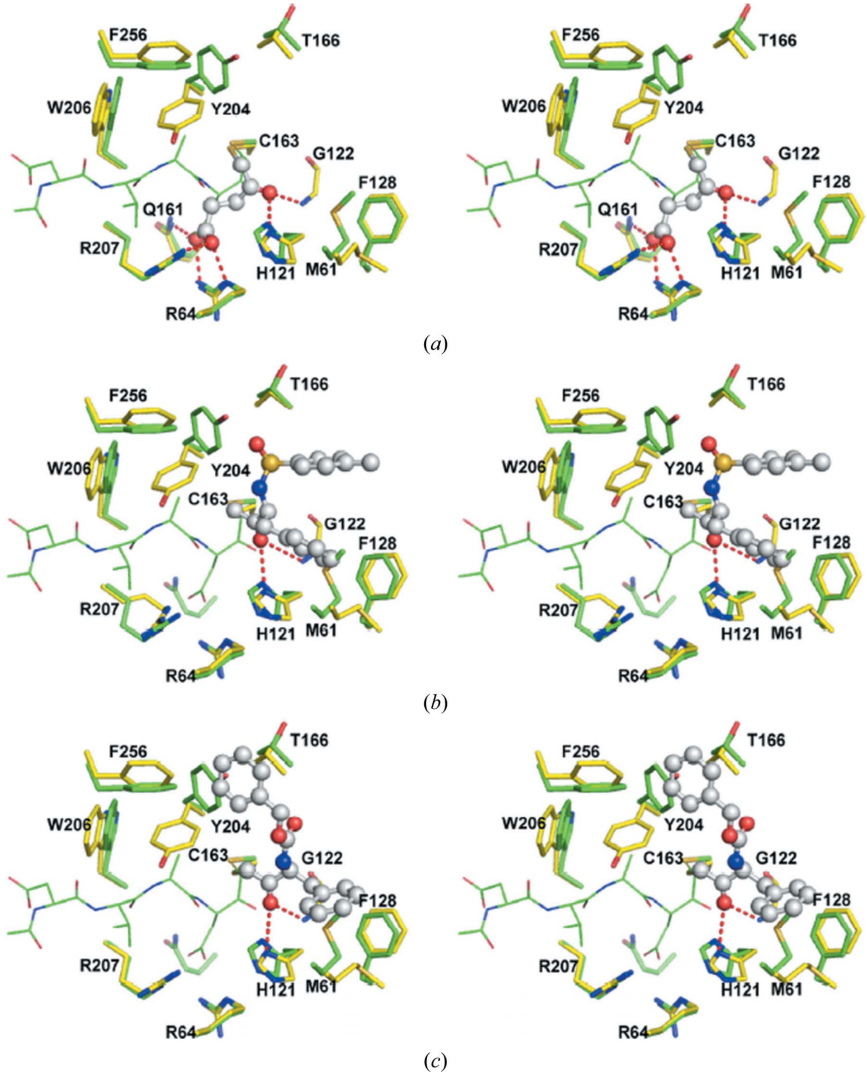


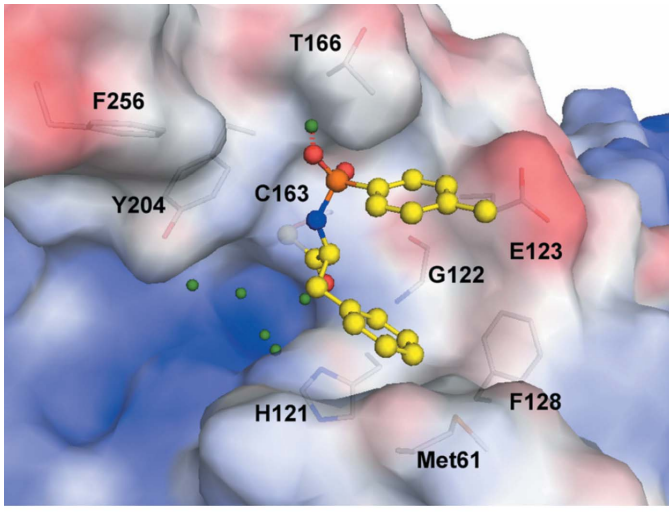
Figure 4
Stereoviews of the active site of caspase-3 (yellow C atoms) in complex with compounds **5**, **1** and **8** (grey C atoms) in (*a*), (*b*) and (*c*), respectively. The structure of the caspase-3–DVAD complex (green C atoms; PDB entry 1cp3) has been superimposed for comparison. A displacement of Tyr204 is observed in all three caspase-3–inhibitor complexes reported here. The binding mode of the carboxylic acid group of compound **5** closely mimics the binding of the carboxylic acid of P1 Asp in known structures and it is involved in ionic interactions with Arg64 and Arg207. Compounds **1** and **8** bind only in the prime site, leaving the S4–S1 sites unoccupied.

a short hydrogen bond to the amide group of Gly122 and a longer hydrogen bond to the side chain of the catalytic His121. This is in contrast to the interaction normally found in the structures of caspase–halomethyl ketone inhibitor complexes and is analogous to the interactions found in the case of caspase–aldehyde inhibitor complexes.

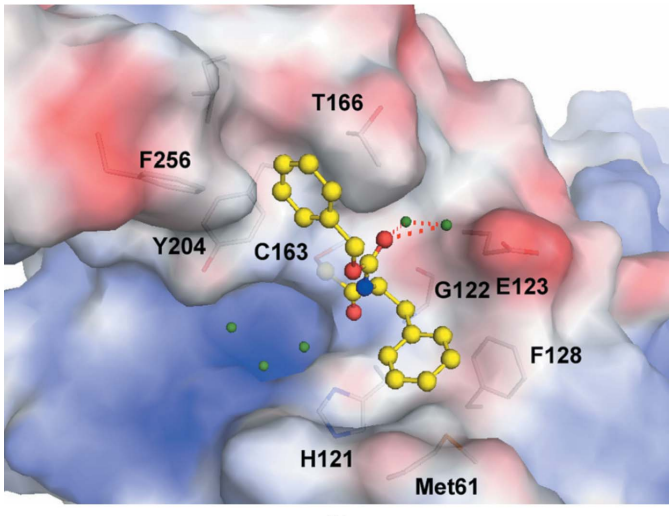
As observed for compounds **1** and **5**, a similar conformational change of Tyr204 is observed for the caspase-3–compound **8** crystal structure. As indicated in Fig. 4(*c*), movement of Tyr204 is warranted to prevent potential steric clashes with one of the phenyl rings of compound **8**. Therefore, the conformation of Tyr204 reported for most of the caspase-3–peptidic inhibitor complexes (Mittl *et al.*, 1997; Rotonda *et al.*, 1996) is caused by the conformation of the P2 Val residue of the inhibitor. Thus, the observed conformational changes of the active-site residues Tyr204 and Arg207, in addition to the prime-site residues Met61 and Phe128, provide a glimpse of the adaptability and plasticity of the substrate-binding site.

The computed binding mode for compound **5** is essentially identical to the binding mode for this compound in the experimental structure. In contrast, the computed binding modes for compounds **1** and **8** are very different from those observed crystallographically. Fig. 6 shows a superposition of the X-ray structures and the computed models of caspase-3 in complex with compound **1** (Fig. 6*a*) and compound **8** (Fig. 6*b*), respectively. There are some prominent differences in the side-chain conformations of active-site amino-acid residues between the crystallographic structures and the protein conformations used as templates for docking. These conformational changes, as well as some crystallographic water molecules that are important for the binding of compounds **1** and **8**, might be responsible for the observed discrepancies between the computed and experimental binding modes.

In the X-ray structure of the caspase-3–compound **1** complex the side chain of Tyr204 occupies the space where the benzyl moiety of the modelled compound **1** binds (Fig. 6*a*). However, the corresponding benzyl group of the ligand in the crystal structure binds to a site that is occupied by the side chain of Met61 in the modelled complex. Two water molecules seem to play an important role in the crystallographically observed binding of compound **1**. The carbonyl

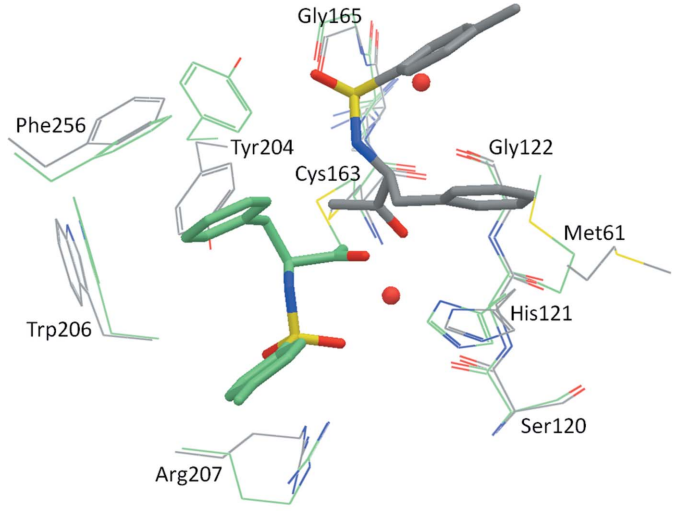


(*a*)

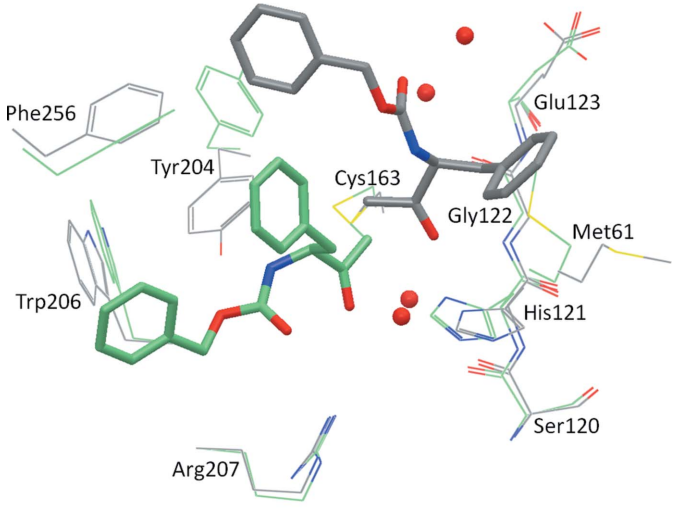


(*b*)

Figure 5
Surface representation illustrating the differences in the binding modes of caspase-3 in complex with compounds **1** and **8** in (*a*) and (*b*), respectively. The unoccupied S1 pocket is populated by water molecules (green spheres). Most of the hydrogen bonds in the prime site are mediated by water molecules. The red and blue rendering indicates negatively and positively charged regions, respectively.



(*a*)



(*b*)

Figure 6
Superposition of the X-ray structure (grey C atoms) and the computed model (green C atoms) of caspase-3 in complex with compounds **1** and **8** in (*a*) and (*b*), respectively. Ligands are shown in bold. The positions of some water molecules that are involved in water-mediated contacts between the ligand and the protein in the respective X-ray structures are shown as red spheres.

group forms water-mediated contacts to the main-chain carbonyl group of Ser120 and to the backbone NH group of Cys163. Additionally, the sulfonyl group is stabilized by a water-mediated contact to the main-chain carbonyl group of Gly122 and the NH group of Gly165. The low temperature factors of the two water molecules indicate that these water molecules are important for the binding of compound **1**. In the modelled complex the benzyl moiety of the docked ligand is stabilized by aromatic and hydrophobic interactions with residues Phe256, Tyr204 and Trp206. The carbonyl group forms a hydrogen bond to the imidazole ring of His121. Furthermore, the sulfonyl group binds to the guanidinium group of Arg207 and the phenyl ring of the attached toluyl group seems to be stabilized by cation–π interactions with Arg207 and forms hydrophobic contacts to some other residues.

In the X-ray structure of the caspase-3–compound **8** complex the benzyl group of the benzyloxycarbonyl group and the phenyl ring of the phenylalanyl moiety of the ligand bind in sites that are occupied by the side chains of Tyr204 and Met61 in the modelled complex, respectively (Fig. 6*b*). Again, water molecules seem to be important for ligand binding. The ketone O atom forms a water-mediated hydrogen-bond network to the main-chain carbonyl group of Ser120, the NH group of Cys163 and the imidazole group of His121. The carbonyl group of the benzyloxycarbonyl moiety is also stabilized by water-mediated contacts to the main-chain carbonyl group of Gly122 and the carboxylate group of Glu123. In contrast, the two phenyl groups of the docked compound **8** are stabilized by hydrophobic contacts to side-chain atoms of Tyr204, Trp206 and Phe256. Moreover, the carbonyl group of the benzyloxycarbonyl moiety seems to form a hydrogen bond to the guanidinium group of Arg207.

In conclusion, ligand-induced conformational changes and *a priori* non-obvious water-molecule positions might be an explanation of the observed discrepancies between the computed and experimental binding modes for compounds **1** and **8**. The X-ray structures revealed that some water molecules play an important role in the binding of these two ligands. The incidence and positions of these water molecules depend upon or can be induced by a particular ligand and therefore cannot easily be accurately predicted. As previously indicated, all crystallographic water molecules were removed from the protein structures prior to docking. Furthermore, the water molecules essential for the binding of ligands **1** and **8** were not conserved in any of the four protein structures (PDB entries 1cp3, 1nme, 1nmq and 1nms) used as templates for the docking experiments. In the X-ray structures of caspase-3 in complex with compounds **1** and **8** the temperature factors of the ligand atoms are clearly higher than those of the atoms of the surrounding protein amino-acid residues, indicating that some of the ligand-binding interactions are not very strong. The weakness of the binding is also reflected in the only moderate inhibitory activities. In the X-ray structures of the four caspase-3 structures used for the redocking and cross-docking experiments no mediating water molecules were essential for ligand binding. These experiments indicate that the docking procedure used seems to give better results in predicting ligand-binding modes for protein–ligand complexes in which water molecules do not play a mediating role (Fig. 1 and Table 1); indeed, our validation experiments demonstrated that redocking (which was successful in all four experiments) worked clearly better than cross-docking (which was successful in six of 12 cases). Therefore, ligand-induced conformational changes as well as the presence and location of water molecules that are important for ligand binding seem to have a significant influence on docking results.

Finally, the high-resolution crystal structures reported here could be used for the design of more potent inhibitors. As an example, one could design derivatives of compounds **1** and **8** which replace the water-mediated hydrogen bonds (Fig. 5) by direct hydrogen-bonding interactions with the protein.

We would like to thank the staff at the Swiss Light Source (PSI, Villigen, Switzerland) for synchrotron beam time. We thank Armin Widmer (Novartis Pharma, Basel) for the software *WITNOTP*. Financial support from the Swiss National Science Foundation and the Baugartenstiftung (Zürich, Switzerland) is gratefully acknowledged. We thank the Drug Synthesis and Chemistry Branch, the Developmental Program and NCI for providing us with chemical compounds. We would also like to thank the anonymous referees for valuable comments and suggestions on the manuscript.

References

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Budin, N., Majeux, N. & Caflisch, A. (2001). *Biol. Chem.* **382**, 1365–1372.
Caflisch, A., Fischer, S. & Karplus, M. (1997). *J. Comput. Chem.* **18**, 723–743.
Cecchini, M., Kolb, P., Majeux, N. & Caflisch, A. (2004). *J. Comput. Chem.* **25**, 412–422.
Choong, I. C., Lew, W., Lee, D., Pham, P., Burdett, M. T., Lam, J. W., Wiesmann, C., Luong, T. N., Fahr, B., DeLano, W. L., McDowell, R. S., Allen, D. A., Erlanson, D. A., Gordon, E. M. & O'Brien, T. (2002). *J. Med. Chem.* **45**, 5005–5022.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Ekici, O. D., Li, Z. Z., Campbell, A. J., James, K. E., Asgian, J. L., Mikolajczyk, J., Salvesen, G. S., Ganesan, R., Jelakovic, S., Grütter, M. G. & Powers, J. C. (2006). *J. Med. Chem.* **49**, 5728–5749.
Ekonomiuk, D., Su, X.-C., Ozawa, K., Bodenreider, C., Lim, S. P., Otting, G., Huang, D. & Caflisch, A. (2009). *J. Med. Chem.* **52**, 4860–4868.
Erlanson, D. A., Lam, J. W., Wiesmann, C., Luong, T. N., Simmons, R. L., DeLano, W. L., Choong, I. C., Burdett, M. T., Flanagan, W. M., Lee, D., Gordon, E. M. & O'Brien, T. (2003). *Nature Biotechnol.* **21**, 308–314.
Fuentes-Prior, P. & Salvesen, G. S. (2004). *Biochem. J.* **384**, 201–232.
Ganesan, R., Jelakovic, S., Campbell, A. J., Li, Z. Z., Asgian, J. L., Powers, J. C. & Grütter, M. G. (2006). *Biochemistry*, **45**, 9059–9067.
Ganesan, R., Mittl, P. R., Jelakovic, S. & Grütter, M. G. (2006). *J. Mol. Biol.* **359**, 1378–1388.
Garcia-Calvo, M., Peterson, E. P., Rasper, D. M., Vaillancourt, J. P., Zamboni, R., Nicholson, D. W. & Thornberry, N. A. (1999). *Cell Death Differ.* **6**, 362–369.
Goode, D. R., Sharma, A. K. & Hergenrother, P. J. (2005). *Org. Lett.* **7**, 3529–3532.
Grütter, M. G. (2000). *Curr. Opin. Struct. Biol.* **10**, 649–655.
Guha, R., Howard, M. T., Hutchison, G. R., Murray-Rust, P., Rzepa, H., Steinbeck, C., Wegner, J. & Willighagen, E. L. (2006). *J. Chem. Inf. Model.* **46**, 991–998.
Hardy, J. A., Lam, J., Nguyen, J. T., O'Brien, T. & Wells, J. A. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 12461–12466.
James, K. E., Asgian, J. L., Li, Z. Z., Ekici, O. D., Rubin, J. R., Mikolajczyk, J., Salvesen, G. S. & Powers, J. C. (2004). *J. Med. Chem.* **47**, 1553–1574.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Keller, N., Grütter, M. G. & Zerbe, O. (2010). *Cell Death Differ.* **17**, 710–718.
Keller, N., Mares, J., Zerbe, O. & Grütter, M. G. (2009). *Structure*, **17**, 438–448.
Kolb, P. & Caflisch, A. (2006). *J. Med. Chem.* **49**, 7384–7392.
MacCorkle, R. A., Freeman, K. W. & Spencer, D. M. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 3655–3660.
Majeux, N., Scarsi, M., Apostolakis, J., Ehrhardt, C. & Caflisch, A. (1999). *Proteins*, **37**, 88–105.
Majeux, N., Scarsi, M. & Caflisch, A. (2001). *Proteins*, **42**, 256–268.
McCluggage, W. G., Ganesan, R., Hirschowitz, L., Miller, K. & Rollason, T. P. (2006). *Am. J. Surg. Pathol.* **30**, 209–215.
Mittl, P. R. E., Di Marco, S., Krebs, J. F., Bai, X., Karanewsky, D. S., Priestle, J. P., Tomaselli, K. J. & Grütter, M. G. (1997). *J. Biol. Chem.* **272**, 6539–6547.
No, K. T., Grant, J. A., Jhon, M. S. & Scheraga, H. A. (1990). *J. Phys. Chem.* **94**, 4740–4746.

No, K. T., Grant, J. A. & Scheraga, H. A. (1990). *J. Phys. Chem.* **94**, 4732–4739.
O'Brien, T. & Lee, D. (2004). *Mini Rev. Med. Chem.* **4**, 153–165.
Piana, S., Sulpizi, M. & Rothlisberger, U. (2003). *Biochemistry*, **42**, 8720–8728.
Rotonda, J., Nicholson, D. W., Fazil, K. M., Gallant, M., Gareau, Y., Labelle, M., Peterson, E. P., Rasper, D. M., Ruel, R., Vaillancourt, J. P., Thornberry, N. A. & Becker, J. W. (1996). *Nature Struct. Biol.* **3**, 619–625.
Scarsi, M., Majeux, N. & Caflisch, A. (1999). *Proteins*, **37**, 565–575.
Schweizer, A., Briand, C. & Grütter, M. G. (2003). *J. Biol. Chem.* **278**, 42441–42447.
Watt, W., Koeplinger, K. A., Mildner, A. M., Heinrikson, R. L., Tomasselli, A. G. & Watenpaugh, K. D. (1999). *Structure*, **7**, 1135–1143.
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. (1994). *Nature (London)*, **370**, 270–275.
Winn, M. D. *et al.* (2010). *Acta Cryst.* D**67**, 235–242.