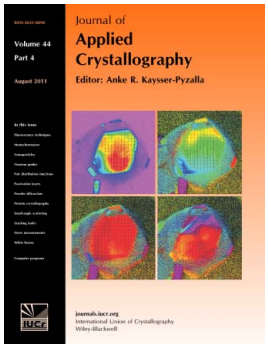
Journal of
Applied Crystallography
ISSN 0021-8898
Editor: **Anke R. Kaysser-Pyzalla**

In-crystal affinity ranking of fragment hit compounds reveals a relationship with their inhibitory activities

Junji Yamane, Min Yao, Yong Zhou, Yasuyuki Hiramatsu, Kenichiro Fujiwara, Tohru Yamaguchi, Hiroto Yamaguchi, Hiroko Togame, Hideki Tsujishita, Hiroshi Takemoto and Isao Tanaka

J. Appl. Cryst. (2011). **44**, 798–804



Journal of Applied Crystallography covers a wide range of crystallographic topics from the viewpoints of both techniques and theory. The journal presents papers on the application of crystallographic techniques and on the related apparatus and computer software. For many years, the *Journal of Applied Crystallography* has been the main vehicle for the publication of small-angle scattering papers and powder diffraction techniques. The journal is the primary place where crystallographic computer program information is published.

Journal of
Applied
Crystallography
ISSN 0021-8898

Received 28 January 2011
Accepted 10 May 2011

In-crystal affinity ranking of fragment hit compounds reveals a relationship with their inhibitory activities

Junji Yamane,[a,b] Min Yao,[a] Yong Zhou,[c] Yasuyuki Hiramatsu,[d] Kenichiro Fujiwara,[d] Tohru Yamaguchi,[d] Hiroto Yamaguchi,[b] Hiroko Togame,[b] Hideki Tsujishita,[d] Hiroshi Takemoto[b] and Isao Tanaka[a]*

[a]Graduate School of Life Science, Hokkaido University, 001-0021 Sapporo, Japan, [b]Shionogi Innovation Center for Drug Discovery, Discovery Research Laboratories, Shionogi & Co. Ltd, 001-0021 Sapporo, Japan, [c]School of Software, Dalian University of Technology, 116620 Dalian, People's Republic of China, and [d]Discovery Research Laboratories, Shionogi & Co. Ltd, 553-0002 Osaka, Japan. Correspondence e-mail: tanaka@castor.sci.hokudai.ac.jp

Fragment-based drug discovery (FBDD), which is a molecular build-up strategy from small scaffolds, has recently become a promising approach for lead-compound generation. Although high-throughput protein crystallography is usually used to determine the protein–ligand complex structure and identify potential hit compounds, the relationship between the quality of the F_o–F_c maps of hit compounds and their inhibitory activities has rarely been examined. To address this issue, crystallographic competition experiments were carried out to determine the relative order of the in-crystal binding affinities using five hit compounds of bovine pancreatic trypsin inhibitors. Soaking experiments of all combinations of the five hit compounds were used to define the in-crystal affinity ranking. Based on characterization by a high-concentration bioassay, a clear correlation was observed between in-crystal binding affinities and the inhibitory activities in solution. Moreover, the correlation analysis revealed that X-ray-based screening can detect a weak hit compound with inhibitory activity below the limit of detection, even in a high-concentration assay. The proposed crystallographic competition method could function as a valuable tool, not only to select a plausible starting scaffold for subsequent synthetic efforts but also to access structure–activity relationships using fragment compounds with a wider detection limit than a biological assay. The crystallographic validation methodology described here will greatly accelerate the hit-to-lead process during fragment-based and structure-based drug design.

1. Introduction

Recent progress in high-throughput protein crystallography (*e.g.* synchrotron radiation facilities, automount robot systems for data collection and a range of software for automated structure determination) has greatly facilitated the drug discovery process (Blundell *et al.*, 2002; Sharff & Jhoti, 2003; Blundell & Patel, 2004; Maclean *et al.*, 2006). X-ray crystallography can be used to elucidate the detailed interactions between the target protein and key small molecules, such as hit compounds obtained from high-throughput screening (HTS) and their follow-up synthetic derivatives. It also provides opportunities to increase the reliability and efficiency of structure-based drug design (Carlson, 2002; Davis *et al.*, 2003; Teague, 2003; Congreve *et al.*, 2005).

The X-ray-based screening of fragment compounds is an extended application of high-throughput protein crystallography for lead-compound generation (Nienaber *et al.*, 2000; Congreve *et al.*, 2003; Hartshorn *et al.*, 2005). Fragment-based approaches use focused compound libraries composed of molecules that are smaller and functionally simpler than drug molecules. Thus, the so-called hit compounds bearing a few functional moieties usually have lower binding affinity and weaker biological activity than those obtained from the HTS approach. However, fragment-based approaches are expected to gain access more easily to larger chemical spaces than the HTS approach. In addition, hit compounds, which have higher ligand efficiency (ΔG_{bind}/heavy atom count) than those obtained by the HTS approach (Hopkins *et al.*, 2004; Abad-Zapatero & Metz, 2005; Abad-Zapatero *et al.*, 2010), are usually more appropriate for lead-compound generation (Barelier *et al.*, 2010). Therefore, once fragment compounds bound to the target protein have been identified in a difference electron-density map (F_o–F_c map) through X-ray crystallography, they will eventually yield better lead compounds by structure-based drug design. The importance and potential of the fragment-based approach to generate lead compounds

have been supported by a number of previous studies (Hajduk & Greer, 2007; Congreve *et al.*, 2008).

There are three main processes in the fragment-based approach: (i) the design of a library of fragment compounds; (ii) the identification of hit compounds that bind to the target protein; and (iii) a follow-up synthetic process aimed at hit-compound optimization for more efficient candidates possessing drug-like properties. X-ray-based crystallographic screening is a key methodology for hit-compound identification, even though it is more time consuming than other biophysical screening approaches, such as nuclear magnetic resonance (NMR; Shuker *et al.*, 1996; Hajduk *et al.*, 1997; Dalvit *et al.*, 2001; Dalvit, 2009) and surface plasmon resonance (SPR; Karlsson *et al.*, 2000; Neumann *et al.*, 2007; Hämäläinen *et al.*, 2008; Danielson, 2009). Every screening approach has inherent features with methodological limitations (*e.g.* throughput, sensitivity and detection limits). Thus, depending on the size of the designed fragment-compound library, compensatory efforts have been applied to identify potential hit compounds for lead-compound generation. The most prominent feature of the X-ray-based approach is the ability to identify hit compounds and determine interactions with target molecules simultaneously. The F_o–F_c maps in X-ray analysis can always address what, where and how hit compounds bind to the target protein. Thus, the structural understanding obtained allows for both easier hit-compound validation for the starting scaffold and follow-up optimization for generating potent lead compounds. In addition, F_o–F_c maps will sometimes clarify how water molecules regulate the formation of protein–ligand complexes. Moreover, the determined crystallographic parameters of the hit compounds, such as *B* factor and occupancy, will provide insight into some of the atomic contributions to complex formation. However, despite its wide applicability and future potential, the relationship between the in-crystal binding affinities of hit compounds found in the F_o–F_c maps and their inhibitory activities in solution are still not well understood.

To assess this issue, we investigated the relative order of binding affinities of hit compounds within the protein crystal (in-crystal affinity ranking) and compared their inhibitory activities in solution. On the basis of the wealth of structural knowledge from biological analysis and the availability of well diffracting crystals, we chose bovine pancreatic trypsin as a template for relationship analysis (Bode & Schwager, 1975; Rauh *et al.*, 2004; Radisky *et al.*, 2006). To obtain several competitive hit compounds, we first carried out X-ray-based screening using a 62-member focused library and identified 15 hit compounds that bound to the active site. The in-crystal affinity ranking was assessed by pairwise crystallographic competition experiments using the five representative hit compounds selected from each scaffold cluster on the basis of structural simplicity. To characterize the inhibitory activity of all the hit compounds, the approximate 50% inhibitory concentration (IC_{50}) values were estimated by fluorescence-based assay. These experiments showed that there was a clear correlation between the in-crystal affinity ranking and the in-solution inhibitory activity.

2. Materials and methods

2.1. Materials

All fragment compounds were obtained commercially and used without further purification. All other chemicals were of analytical or reagent grade. Bovine pancreatic trypsin and benzoyl-arginine-4-methylcoumaryl-7-amide (Bz-Arg-MCA) were purchased from MP Biomedicals (Solon, Ohio, USA). All crystallization reagents were purchased from Hampton Research (Oklahoma City, Oklahoma, USA). CrystalQuick sitting-drop 96-well plates and small-volume black flat-bottomed 384-well plates were purchased from Greiner-Bio-One GmbH (Solingen, Germany).

2.2. Crystallization

Bovine pancreatic trypsin was crystallized by the sitting-drop method by mixing 1 µl of protein solution (30 mg ml^{-1}) and 1 µl of reservoir solution. The protein was dissolved in a buffer containing 25 m*M* HEPES pH 7.0 and 5 m*M* $CaCl_2$. The BCA assay method (Pierce, Rockford, Illinois, USA) was used for the determination of the protein concentration. Initial crystal screenings were performed using commercially available kits – Crystal Screen and Index Screen (Hampton Research), and Wizard Screen (Emerald BioSystems, Bainbridge Island, Washington, USA). Crystals of the apo form (apo crystals) were grown at 277 K after two weeks with a reservoir solution of 30% polyethylene glycol 3350, 0.2 *M* lithium sulfate and 0.1 *M* Tris–HCl pH 8.5. The crystals obtained were of a suitable size, morphology and diffraction quality for high-resolution crystallographic study.

2.3. Crystallographic screening

To obtain trypsin crystals complexed with each fragment compound, soaking experiments were carried out by transferring apo crystals to sitting drops (5 µl) containing 30% (*w*/*v*) polyethylene glycol 3350, 0.2 *M* lithium sulfate, 10% (*v*/*v*) dimethylsulfoxide (DMSO) and 0.1 *M* Tris–HCl pH 8.5, supplemented with 6 m*M* of the fragment compound. After incubation at 293 K for 10 min under freely diffusing conditions, we obtained a set of crystals complexed with the fragment compounds.

2.4. Crystallographic competition experiment

To investigate the relative order of the binding affinities in the protein crystals (in-crystal affinity ranking), soaking experiments using mixed solutions of two hit compounds were carried out by transferring apo crystals to drops (5 µl) containing 30% (*w*/*v*) polyethylene glycol 3350, 0.2 *M* lithium sulfate, 10% (*v*/*v*) DMSO and 0.1 *M* Tris–HCl pH 8.5, supplemented with 6 m*M* of either fragment compound. After incubation at 293 K for 15 min under freely diffusing conditions, crystals complexed with the fragment compounds were obtained.

2.5. X-ray data collection, processing and structure determination

Prior to data collection, the complexed crystals were transferred to various cryosolutions containing a soaking

solution with 10% ethylene glycol. These were then looped out and frozen in liquid nitrogen for data collection. Data were collected on a Rigaku R-AXIS IV++ imaging-plate area-detector system mounted on a Rigaku MicroMax007 rotating-anode X-ray generator using Cu $K\alpha$ radiation (40 kV, 20 mA). All data sets were indexed, integrated, scaled and merged using the *HKL2000* program package (Otwinowski & Minor, 1997). The statistics are summarized in Supplementary Tables S1–S2.[1]

All the trypsin crystals belonged to the $P2_12_12_1$ space group with a single molecule in the asymmetric unit. The phases were obtained by molecular replacement using the structure complexed with benzamidine as a search model [PDB (Protein Data Bank; Berman *et al.*, 2000) code 1s0r. After the initial phase determination, automatic restrained positional and isotropic *B*-factor refinement was performed using *LAFIRE* (Yao *et al.*, 2006; Zhou *et al.*, 2006) with *REFMAC* (Collaborative Computational Project, Number 4, 1994; Murshudov *et al.*, 1997). In addition, all the binding conformations of the fragment compounds were built and refined using the automatic ligand-detecting and ligand-fitting procedure of *LAFIRE* (data not shown). The final refinement with the ligand and water molecules was carried out using *REFMAC* and *COOT* (Emsley & Cowtan, 2004). All refinement statistics are shown in Supplementary Tables S1–S2. All figures were generated using *PyMol* (DeLano, 2002).

The coordinates and structure factors of the described structures have been deposited in the PDB with accession codes 3rxa (F01), 3rxb (F02), 3rxc (F03), 3rxd (F04), 3rxe (F05), 3rxf (F06), 3rxg (F07), 3rxh (A02), 3rxi (A06), 3rxj (A18), 3rxk (A51), 3rxl (A52), 3rxm (A53), 3rxo (A56), 3rxp (A60), 3ati (F04, F01–F04 cocktail), 3rxq (F05, F01–F05 cocktail), 3atk (F01, F01–A06 cocktail), 3rxr (F01, F01–F03 cocktail), 3atl (F05, F04–F05 cocktail), 3rxs (F04, F04–A06 cocktail), 3rxt (F04, F04–F03 cocktail), 3rxu (F05, F05–A06 cocktail), 3rxv (F05, F05–F03 cocktail) and 3atm (A06, A06–F03 cocktail).

2.6. Fluorescence-based inhibition assay

Trypsin inhibitory activity assays were carried out in 384-well plates at room temperature by measurement of the 7-amino-4-methylcoumarin released from Bz-Arg-MCA, which is cleaved by bovine pancreatic trypsin to release the fluorescent aminomethylcoumarin. The fluorescence was measured at 460 nm using a plate reader (EnVision Multilabel Plate Reader; PerkinElmer, Wellesley, Massachusetts, USA). To determine the IC_{50} values, the reaction mixtures were prepared by the addition of 5 µl of DMSO solution of the compound, 20 µl of enzyme buffer and 25 µl of substrate buffer. The protein concentration of the enzyme buffer was fixed at 400 ng ml^{-1}. The linearity of the fluorescence enhancement over 60 min was confirmed. The substrate concentration was fixed at 100 µ*M* and those of the compounds were varied. The final assay buffer contained 150 m*M* NaCl, 1 m*M* $CaCl_2$, 0.1 mg ml^{-1} BSA (bovine serum albumin) and 50 m*M* Tris–HCl pH 8.0. All kinetic parameters and confidence limits were determined by nonlinear regression analysis using *GraphPad Prism* (GraphPad Software, 2003). The IC_{50} values were calculated using the built-in procedure for sigmoidal dose–response.

3. Results

3.1. Preparation of the focused fragment-compound library

The focused fragment-compound library contained 62 molecules possessing simple scaffolds and functional groups (*e.g.* primary amino, secondary amino, hydroxyl and carboxyl groups). In order to validate the crystallographic experiments, the designed library included seven fragment compounds (F01, F02, F03, F04, F05, F06 and F07) which are known to bind to trypsin, forming hydrogen bonds with Asp189 at the bottom surface of the S1 pocket. The constructed library is illustrated in Supplementary Fig. S1.

3.2. Crystallographic screening

To ascertain whether the apo crystals were suitable for soaking experiments, we first determined the crystal structure

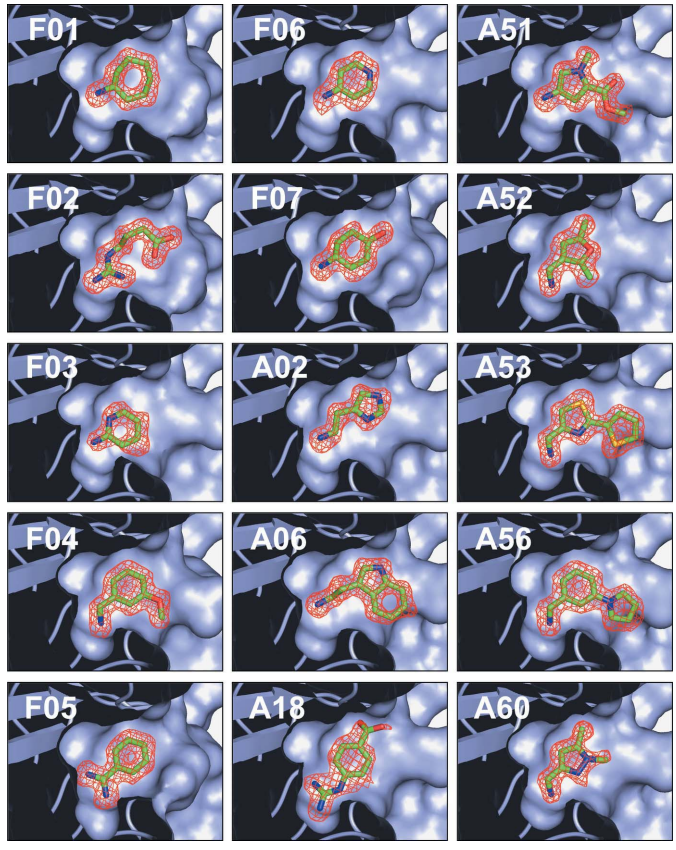


Figure 1
The 15 hit compounds were identified by X-ray-based crystallographic screening using a 62-member focused library. Fragment compounds (green carbon sticks) are superimposed on the σ_A-weighted F_o–F_c map (red mesh, contoured at 2.5σ). For crystallographic data, see Supplementary Table S1.

[1] Supplementary materials discussed in this paper are available from the IUCr electronic archives (Reference: HE5523). Services for accessing these data are described at the back of the journal.

of apo trypsin by molecular replacement. Based on the high-resolution structure of the apo form, we confirmed that no ligand was bound to Asp189 in the F_o–F_c maps. In addition, based on packing analysis, we confirmed that the active site was accessible to the fragment compounds.

In order to obtain a series of competitive hit compounds, we carried out crystallographic screening by soaking experiments using the 62 fragment compounds. After automatic ligand detection and fitting using *LAFIRE*, we identified 15 hit compounds, all of which possessed an amino group bound to Asp189 in the shallow S1 pocket (Fig. 1). All the residual F_o–F_c maps were sufficiently clear to characterize the binding scheme of all the hit compounds. Although they showed some differences in their chemical structures within the context of their scaffolds and functional groups, there was little structural deviation upon complex formation, except the side chain of Gln192 at the entrance to the S1 pocket. As expected, we confirmed the seven known binders (F01, F02, F03, F04, F05, F06 and F07) bound to the S1 pocket, as reported previously. In addition, we found that eight additional fragment compounds (A02, A06, A18, A51, A52, A53, A56 and A60) containing aromatic or aliphatic scaffolds bore the amino group bound to the S1 pocket *via* hydrogen bonds with Asp189. It should be noted that no fragment compound was identified in F_o–F_c maps contoured at 2.5σ.

The amino groups of each hit compound were capable of making a significant contribution to the overall binding energy because of their ability to form hydrogen bonds with the carboxylate group of Asp189 (Fig. 1). In addition, each hydrophobic group could provide additional binding energy for complex formation *via* van der Waals interactions, which are dependent upon shape complementarity.

There are some differences in the hydrogen-bond networks depending on the types of amino group. In the case of an amidine or guanidine moiety (F02, F05, A18), two-on-two hydrogen bonds are formed with the carboxylate group of Asp189. Additional hydrogen bonds are formed with the carbonyl group of Gly214 and with the hydroxyl group of Ser190 *via* a bridging water molecule. In the case of the primary amino group (F01, F03, F04, F06, F07, A02, A06, A51, A52, A53, A56, A60), the NH_2 group of each hit compound forms two hydrogen bonds with the carboxylate group of Asp189, and a bidirectional hydrogen bond with the carbonyl groups of Gly214 and the hydroxyl group of Ser190. Moreover, it forms hydrogen bonds with the carbonyl group of Val223/Trp211 and the amide group of Val223 *via* a bridging water molecule.

3.3. Structural classification of hit compounds

To examine the molecular similarity of the 15 hit compounds, we carried out a structural classification by scaffold-based clustering analysis. Based on the amino-group types and chemical features of the main scaffold, the compounds were categorized into five classes, as illustrated in Fig. 2. Each of the scaffolds I–V contains an analogous series

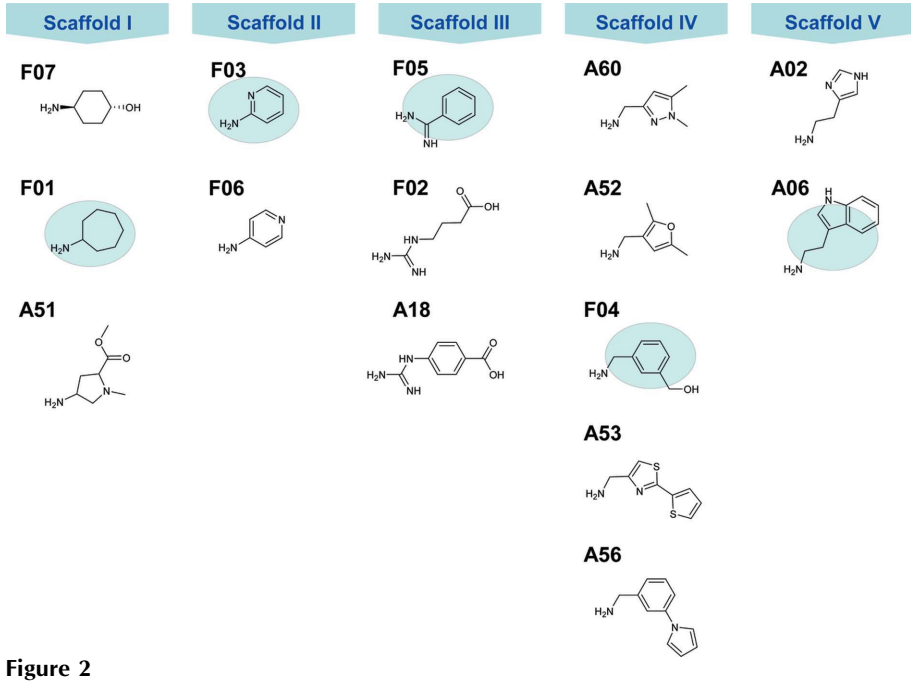


Figure 2
Based on the amino group types and chemical features of the main scaffolds, 15 hit compounds were categorized into five scaffolds, I–V. The five fragment compounds marked with a shaded background were used for crystallographic competition experiments to examine the relative order of their binding affinities *via* crystallography.

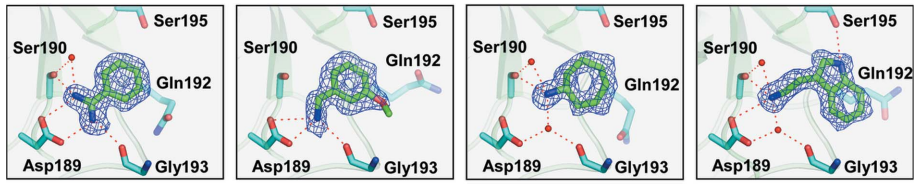


Figure 3
The σ-weighted F_o–F_c maps, contoured at 2.5σ (blue mesh), superimposed with the fragment compounds (green carbon sticks). Interacting side chains and water molecules are shown as sticks and red spheres, respectively. Hydrogen bonds are shown as broken red lines. (*a*) F05 was selected from a mixed solution containing F04 and F05. (*b*) F04 was selected from a mixed solution containing F01 and F04. (*c*) F01 was selected from a mixed solution containing F01 and A06. (*d*) A06 was selected from a mixed solution containing F03 and A06. All F_o–F_c maps of the crystallographic competition experiments are shown in Supplementary Fig. S2. For crystallographic data, see Supplementary Table S2.

of the particular scaffold. Scaffold I contains three hit compounds bearing the five- or six-membered aliphatic ring with the primary amino group. Scaffold II contains two hit compounds bearing the six-membered aromatic ring with the primary amino group (aminopyridine derivatives). Scaffold III contains three hit compounds bearing the aromatic ring with the amidino or guanidine group, and also the aliphatic chain with the guanidine group. Scaffold IV contains five hit compounds bearing the five- or six-membered aromatic ring with the aminomethyl group. Scaffold V contains two hit compounds bearing the five- or six-membered aromatic ring with the aminoethyl group. This classification suggests that the spatial arrangement between the amino group and the hydrophobic group is the consensus structural element of the hit compounds. In addition, the scaffold diversity indicates that crystallographic screening can access a large chemical space with small numbers of compounds, as reported previously (Hopkins *et al.*, 2004; Abad-Zapatero & Metz, 2005; Abad-Zapatero *et al.*, 2010).

3.4. Crystallographic competition experiment

To investigate the relationship between the binding affinities in the protein crystal and the inhibitory activities in solution, we carried out a crystallographic competition experiment using the five representative hit compounds (F01, F03, F04, F05, A06), which were selected from each scaffold class on the basis of structural simplicity. For this purpose, we prepared ten mixed solutions of all possible combinatorial variations containing two hit compounds. The soaking and X-ray experiments were carried out and the F_o–F_c maps were examined.

The results of the crystallographic competition experiments confirmed that clear F_o–F_c maps of the singlet fragment compound were observed in all ten cases. After all pairwise examinations, we determined the relative order of the binding affinities in the protein crystals (in-crystal affinity ranking) of the five representative hit compounds without any contradictions (Figs. 3 and 4, and Supplementary Fig. S2).

3.5. Kinetic study using fluorescence-based substrate

The results of enzyme kinetic studies confirmed that X-ray-based fragment-compound screening can identify various competitive inhibitors, the IC_{50} values of which were calculated to range from 17.2 μ*M* to 8.3 m*M* (Table 1). In addition, we confirmed that subtle scaffold differences in the fragment compounds would lead to large differences in inhibitory activity, even though their manners of binding were almost the same. To examine the ligand efficiency, the binding efficiency index (BEI) values were calculated for each hit compound (Hopkins *et al.*, 2004; Abad-Zapatero & Metz, 2005; Abad-Zapatero *et al.*, 2010). Moreover, X-ray-based fragment-compound screening was able to detect a weak hit compound (F03) with inhibitory activity below the limit of detection, even in a high-concentration assay.

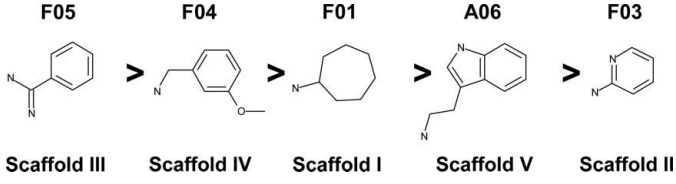


Figure 4
The relative order of the binding affinities in the protein crystals (in-crystal affinity ranking).

Table 1
Inhibitory activities, Hill slope and binding efficiency index (BEI) values against bovine pancreatic trypsin.

For dose–response curve data, see Supplementary Fig. S3. n.d. indicates that the value was not determined.

Hit compound	MW (Da)	IC_{50} (μM)†	Hill slope†	BEI‡	*c* log *P*§
F01	113	1400	−0.83	25.3	1.7
F02	145	2200	−0.97	18.3	−0.67
F03	94	n.d.	n.d.	n.d.	0.66
F04	137	188	−1.1	27.2	1.2
F05	120	17.2	−1.0	39.7	0.97
F06	94	2500	−0.98	27.7	0.66
F07	115	7100	−1.1	18.7	0.25
A02	111	15700	−0.64	16.3	−0.089
A06	160	873	−1.1	19.1	1.7
A18	179	873	−1.1	17.1	0.69
A51	158	2100	−1.1	17.0	−0.81
A52	125	2400	−0.96	21.0	1.6
A53	196	232	−0.79	19.0	2.6
A56	172	83.3	−0.89	24.0	2.2
A60	125	8300	−0.97	16.7	0.81

† IC_{50} and Hill slope values were calculated using the built-in procedure for sigmoidal dose–response. ‡ BEI (Abad-Zapatero & Metz, 2005) calculated for trypsin inhibition, BEI = $-\log(IC_{50})/M$, where M is the mass of the compound in kDa. § *c* log *P* values were calculated using the *Molecular Operating Environment* (*MOE*) program (Chemical Computing Group, 2010).

3.6. Relationship between in-crystal binding affinities and their inhibitory activities

The in-crystal affinity ranking of the five hit compounds and their inhibitory activities showed that there was a clear positive correlation, except for one apparent reversal. According to the IC_{50} values, A06 (IC_{50} = 873 μ*M*) was a more potent inhibitor than F01 (IC_{50} = 1.4 m*M*). However, the crystallographic competition experiment based on the F_o–F_c map indicated that F01 was bound rather than A06.

There are several possible explanations for this outlier. The first is that the weak protein–ligand interaction in the protein crystal may be affected by the packing effect. The restricted S1 pocket may incorporate a more rigid and smaller hit compound (F01). While the NH_2 group of F01 (Scaffold I) is bonded directly to the cycloheptane moiety, the NH_2 group of A06 (Scaffold IV) is connected to the indole moiety with a flexible ethylene linker.

The second possibility is that the partial affinity of F01 in the S1 pocket is higher than that of A06. This interpretation is consistent with the order of ligand efficiency estimated from the IC_{50} values (Table 1). In the competitive experiment, the more efficient binder (F01, BEI = 25.3) would occupy its binding site rather than the less efficient binder (A06, BEI = 19.1).

The third and the most likely possibility is related to the confidence limit of the high-concentration assay using the fluorescence-based substrate (Bz-Arg-MCA), because there are some difficulties in determining the IC_{50} values of weak inhibitors. The results of the high-concentration assay confirmed the dose–response curve characteristics of simple Michaelis–Menten kinetics. In addition, each IC_{50} value was validated by kinetic parameters. However, we observed 80% inhibitory activities for F01 and A06, although the maximum ligand conditions were adjusted to 6 m*M*. Therefore, the IC_{50} values of weak hit compounds, such as F01 and A06, may include some unexpected experimental errors (*e.g.* inductive measurement errors and/or nonspecific interactions due to high compound concentration). Thus, the small differences in IC_{50} values between A06 (IC_{50} = 873 μ*M*) and F01 (IC_{50} = 1.4 m*M*) were not sufficient to examine the order of the weak binding affinity.

4. Discussion

The ligand efficiency (ΔG_{bind}/heavy atom count), which is simply the binding free energy proposed by Hopkins *et al.* (2004), provides insight into the characterization of hit compounds and their follow-up derivatives. However, some experimental obstacles, such as errors of measurement arising from compound effects (self-fluorescence, fluorescence quenching, self-aggregation *etc.*) under high-concentration conditions, would prevent the assessment of the binding affinities and inhibitory activities of weak hit compounds. Moreover, some key problems in fragment-compound screening by NMR, SPR and high-concentration assay are the prevalence of false-positive, false-negative and nonspecific (promiscuous) inhibition. Therefore, another practical validation method would be required for proper hit-compound characterization (Rishton, 2003; Seidler *et al.*, 2003; Shoichet, 2006). One important practical advantage of fragment-compound screening by X-ray diffraction would be its ability to detect only selective binders. If the protein–ligand complex crystals diffract well enough for structure determination, the electron-density maps of non-selective binders are never revealed in the F_o–F_c maps.

In this study, we have determined for the first time the in-crystal affinity ranking of hit compounds by a competition approach with cocktail solutions containing two hit compounds, and shown that it is closely correlated with the IC_{50} values of these compounds. To understand the effect of the dynamic disorder induced by thermal motion on the in-crystal affinity ranking, we examined their F_o–F_c maps and *B* factors. However, there were no significant differences in these crystallographic data. Also, it is well known that each hydrophobic group of the hit compounds usually causes an increase in the binding affinity to the target protein. Thus, to characterize the effect of the hydrophobic properties of fragment hit compounds on their in-crystal affinity ranking, we analysed the relative order of their in-crystal binding affinities and calculated log octanol/water partition coefficients ($c\log P$). Although we designed an ideal situation for studying protein–ligand complex formation without a considerable induced-fit effect (rigid-body manner), there was no significant correlation between experimental in-crystal scores and computational values. This finding would suggest a difficulty in the computational examination and prediction of the weak interactions between proteins and ligands such as fragment hit compounds. In contrast with high-affinity complex formation, some weak water-mediated effects (*e.g.* solvation and desolvation effects of the hit compound and its binding site within the protein, the structural rearrangement effect of water molecules located in the protein and hit-compound interface, and so on) might need to be taken into account when considering the weak interaction between the protein and the fragment hit compound.

5. Conclusions

The described case study demonstrates the successful assessment of in-crystal affinity ranking through crystallographic competition experiments. After hit-compound identification by X-ray-based screening, it provides an easy way of examining the relative order of the in-crystal binding affinity, because only optimization of the soaking conditions would be needed. We believe that the proposed method using hit compounds can be used not only to investigate the relative order of binding affinity in protein crystals (in-crystal affinity ranking) for hit-compound validation, but also to validate potential hit compounds and follow-up synthetic derivatives for lead-compound generation. In addition, this method would be able to access structure–activity relationships using various fragment compounds with a wider affinity range than a biological assay, even when no inhibition activity is observed. Therefore, fragment-based drug design combined with in-crystal affinity ranking, which is a direct readout from protein–ligand complex formation, could be one of the structure-guided solutions for a hit-to-lead strategy.

At present, however, the absolute generality of in-crystal affinity rankings is still unknown. In the case of trypsin, no major conformational change normally occurs within the active site. In addition, all of the fragment compounds used in this study possessed suitable chemical and physical properties for soaking experiments. Therefore, this trypsin study is considered to be an ideal situation for determining in-crystal affinity rankings. Unlike trypsin, some other proteins may change their conformation during protein–ligand complex formation. Therefore, further studies are needed to investigate how the in-crystal affinity ranking is influenced by conformational change. It is well known that some small molecules, called 'crystal crackers', will cause serious physical damage to protein crystals (*e.g.* cracking, increasing mosaicity *etc.*), even if they are potent inhibitors (Danley, 2006; Skarzynski & Thorpe, 2006). The lattice size of the crystal or the packing of the protein molecules may sometimes be among the most important factors.

In this crystallographic competition experiment, all hit compounds were identified as singlets in F_o–F_c maps. However, in some cases, the F_o–F_c maps may indicate simul-

taneous binding of constituent compounds. Further studies are currently underway in our laboratory to determine the individual occupancy depending on the F_o–F_c map.

The authors thank H. Arita, H. Kondo, K. Yasuo, K. Hattori and Y. Sakou (Shionogi & Co. Ltd) for valuable scientific discussions. This work was supported by funding from the Ministry of Education, Culture, Science, Sports and Technology of Japan for the 'Innovation COE Program for Future Drug Discovery and Medical Care'.

References

Abad-Zapatero, C. & Metz, J. T. (2005). *Drug Discov. Today*, **10**, 464–469.
Abad-Zapatero, C., Perišić, O., Wass, J., Bento, A. P., Overington, J., Al-Lazikani, B. & Johnson, M. E. (2010). *Drug Discov. Today*, **15**, 804–811.
Barelier, S., Pons, J., Gehring, K., Lancelin, J. M. & Krimm, I. (2010). *J. Med. Chem.* **53**, 5256–5266.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalow, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Blundell, T. L., Jhoti, H. & Abell, C. (2002). *Nat. Rev. Drug Discov.* **1**, 45–54.
Blundell, T. L. & Patel, S. (2004). *Curr. Opin. Pharmacol.* **4**, 490–496.
Bode, W. & Schwager, P. (1975). *J. Mol. Biol.* **98**, 693–717.
Carlson, H. A. (2002). *Curr. Opin. Chem. Biol.* **6**, 447–452.
Chemical Computing Group (1996). *MOE*. Version 2010.10. Chemical Computing Group Inc., Montreal, Canada.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Congreve, M., Carr, R., Murray, C. & Jhoti, H. (2003). *Drug Discov. Today*, **8**, 876–877.
Congreve, M., Chessari, G., Tisi, D. & Woodhead, A. J. (2008). *J. Med. Chem.* **51**, 3661–3680.
Congreve, M., Murray, C. W. & Blundell, T. L. (2005). *Drug Discov. Today*, **10**, 895–907.
Dalvit, C. (2009). *Drug Discov. Today*, **14**, 1051–1057.
Dalvit, C., Fogliatto, G., Stewart, A., Veronesi, M. & Stockman, B. (2001). *J. Biomol. NMR*, **21**, 349–359.
Danielson, U. H. (2009). *Curr. Top. Med. Chem.* **9**, 1725–1735.
Danley, D. E. (2006). *Acta Cryst.* D**62**, 569–575.
Davis, A. M., Teague, S. J. & Kleywegt, G. J. (2003). *Angew. Chem. Int. Ed.* **42**, 2718–2736.
DeLano, W. L. (2002). *pyMOL*. DeLano Scientific LLC, San Carlos, California, USA, http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
GraphPad Software (2003). *GraphPad Prism*. Version 4. GraphPad Software Inc., La Jolla, California, USA.
Hajduk, P. J. & Greer, J. (2007). *Nat. Rev. Drug Discov.* **6**, 211–219.
Hajduk, P. J., Olejniczak, E. T. & Fesik, S. W. (1997). *J. Am. Chem. Soc.* **119**, 12257–12261.
Hämäläinen, M. D., Zhukov, A., Ivarsson, M., Fex, T., Gottfries, J., Karlsson, R. & Björsne, M. (2008). *J. Biomol. Screen.* **13**, 202–209.
Hartshorn, M. J., Murray, C. W., Cleasby, A., Frederickson, M., Tickle, I. J. & Jhoti, H. (2005). *J. Med. Chem.* **48**, 403–413.
Hopkins, A. L., Groom, C. R. & Alex, A. (2004). *Drug Discov. Today*, **9**, 430–431.
Karlsson, R., Kullman-Magnusson, M., Hämäläinen, M. D., Remaeus, A., Andersson, K., Borg, P., Gyzander, E. & Deinum, J. (2000). *Anal. Biochem.* **278**, 1–13.
Maclean, E. J., Rizkallah, P. J. & Helliwell, J. R. (2006). *Eur. Pharm. Rev.* **2**, 71–76.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Neumann, T., Junker, H. D., Schmidt, K. & Sekul, R. (2007). *Curr. Top. Med. Chem.* **7**, 1630–1642.
Nienaber, V. L., Richardson, P. L., Klighofer, V., Bouska, J. J., Giranda, V. L. & Greer, J. (2000). *Nat. Biotechnol.* **18**, 1105–1108.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Radisky, E. S., Lee, J. M., Lu, C. J. & Koshland, D. E. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 6835–6840.
Rauh, D., Klebe, G. & Stubbs, M. T. (2004). *J. Mol. Biol.* **335**, 1325–1341.
Rishton, G. M. (2003). *Drug Discov. Today*, **8**, 86–96.
Seidler, J., McGovern, S. L., Doman, T. N. & Shoichet, B. K. (2003). *J. Med. Chem.* **46**, 4477–4486.
Sharff, A. & Jhoti, H. (2003). *Curr. Opin. Chem. Biol.* **7**, 340–345.
Shoichet, B. K. (2006). *J. Med. Chem.* **49**, 7274–7277.
Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. (1996). *Science*, **274**, 1531–1534.
Skarzynski, T. & Thorpe, J. (2006). *Acta Cryst.* D**62**, 102–107.
Teague, S. J. (2003). *Nat. Rev. Drug Discov.* **2**, 527–541.
Yao, M., Zhou, Y. & Tanaka, I. (2006). *Acta Cryst.* D**62**, 189–196.
Zhou, Y., Yao, M. & Tanaka, I. (2006). *J. Appl. Cryst.* **39**, 57–63.