ChemComm

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: S. Fu, Q. Nie, Y. Ma, P. Song, X. Ren, C. Luo, L. Shang and Z. Yin, *Chem. Commun.*, 2018, DOI: 10.1039/C8CC00469B.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemcomm

View Article Online
DOI: 10.1039/C8CC00469B

Journal Name

COMMUNICATION

Target-guided Screening of Fragments (TGSOF) in the Discovery of Inhibitors against EV-A71 3C Protease

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Shengjun Fu,[a] Quandeng Nie,[a] Yuying Ma,[a] Ping Song,[a] Xuejiao Ren,[a] Cheng Luo,[a] Luqing Shang,[*,a] and Zheng Yin[*,a]

Target-guided screening of fragments (TGSOF) was developed and employed in the identification of EV-A71 3C protease ($3C^{pro}$) inhibitors. We identified 4-acetylpyridine and 3-acetylpyridine as effective P3 fragments of an inhibitor and obtained the corresponding irreversible inhibitors 12c and 12f via this method. Furthermore, based on 12c and 12f, we have obtained reversible inhibitors 17c and 17f. These results demonstrated that TGSOF is a useful strategy for identifying suitable fragments in developing leads in drug discovery.

Human enterovirus A71 (EV-A71) is a major etiologic agent of hand, foot and mouth disease (HFMD).[1] Several large outbreaks of HFMD have occurred in children since 1997, mostly in the Asia-Pacific region.[2] Several approaches have been attempted to prevent enterovirus infections, but no effective antiviral agents are available for the clinical treatment of EV-A71.[1] EV-A71, a member of the genus *Enterovirus* in the family *Picornaviridae*, contains a single-stranded, positive-sense RNA genome that encodes a large poly-protein precursor.[2] The poly-protein is cleaved into four structural proteins (VP1-VP4) and seven non-structural proteins (2A-2C and 3A-3D).[2] $3C^{pro}$ is required for processing all the cleavage sites on the polyprotein, except VP1/2A and 3C/3D. The critical roles that it plays for the viral replication make $3C^{pro}$ an attractive drug target for EV-A71 infection. Therefore, the critical role of the $3C^{pro}$ makes it an attractive target for anti-EV-A71 agent discovery.[1]

With the aim of developing antivirals for the treatment of EV-A71, several efforts have been made to discover inhibitors of $3C^{pro}$ using traditional medicinal chemistry methods.[1, 3] In a typical and traditional process, biologically active compounds need to be synthesized prior to being assessed for activity.[4] This is a burdensome process as organic synthesis often is the most time-consuming part of a medicinal chemistry project. In the last decade, "target-guided synthesis (TGS)", "tethering in situ click chemistry (TISCC)" and "templated assembly of protein-binding fragments" have demonstrated great promise in streamlining drug discovery processes by combining screening and synthesis into a single step, where the biological target molecule is used as a template for the construction of its own inhibitor through self-assembly within the active site of the target.[4-5] These strategies are based on the notion that when the reaction product binds tightly with the target molecule, the reaction will be driven by the 3D environment of the target molecule.[6] Inspired by these methods, we present a related approach, i.e., target-guided screening of fragments (TGSOF) (Figure 1). Similar to those previously described methods, if a reactive substrate binds tightly to the active pocket of the biomolecule, the corresponding screening reaction in the pocket could be promoted by the 3D environment of the biomolecule. Therefore, TGSOF is proposed for screening drug fragments and obtaining lead compounds. In TGSOF, according to the screening of different fragments of the inhibitor, the corresponding ligation building block (LBB) is designed and synthesized. A functional group can be introduced at a specific site of the target biomolecule using a LBB. Therefore, compared to TGS, TISCC, and "templated assembly of protein-binding fragments", it is ensured that the screening reaction in the TGSOF can occur in a designated location. Furthermore, the screening efficiency was proved and a certain fragment or region of an inhibitor was precisely found. TGSOF can rapidly provide effective fragments for drug discovery and highly active inhibitors are also more likely to be obtained by assembling the corresponding fragment and LBB.

Herein, TGSOF was employed to screen the P3 fragment of a

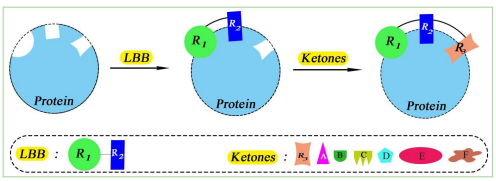


Figure 1.The TGSOF strategy. First, a reactive functional group is site-specifically introduced to the active pocket of the target protein using the ligation building block (LBB). Second, fragment molecules are screened by the modified protein with the LBB. Finally, potential inhibitors are assembled.

[a] *College of Pharmacy, State Key Laboratory of Medicinal Chemical Biology and Tianjin Key Laboratory of Molecular Drug Research, Nankai University, Tianjin 300071, China. E-mail: shanglq@nankai.edu.cn, zheng_yin@nankai.edu.cn*

†*Electronic Supplementary Information (ESI) available: Experimental procedures, characterization data, and additional spectra. See DOI: 10.1039/x0xx00000x*

View Article Online
DOI: 10.1039/C8CC00469B

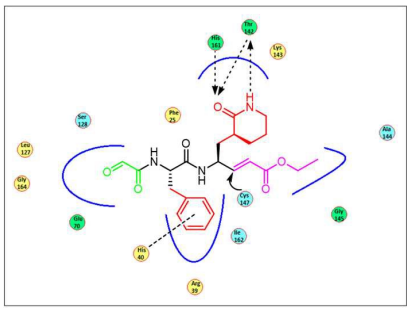


Figure 2. Schematic diagram of the LBB（**8**） interacting with EV-A71 3C^{pro}. The (S)-δ-lactam ring and the benzene group were required to ensure specificity for the 3C^{pro}. The α,β-unsaturated ester was used as the linker to 3C^{pro}. The α-ketoaldehyde was used for the reaction screening.

3C^{pro} inhibitor. Suitable P3 fragments have been identified, leading to several highly active inhibitors. The positive results obtained demonstrate the feasibility of TGSOF.

With the aim of selecting the P3 fragment of the EV-A71 3C^{pro} inhibitor, a LBB was designed based on the following requirements (Figure 2). First, as Cys147 plays a critical role in the catalytic site of EV-A71 3C^{pro}, a covalent-irreversible group, α,β-unsaturated ester, was introduced to obtain strong binding of the LBB with the active site of 3C^{pro}. Second, the P1 and P2 substituents of EV-A71 3C^{pro} inhibitors should preferentially feature a (S)-δ-lactam ring and a benzyl group, respectively, to efficiently occupy the catalytic pockets of the EV-A71 3C^{pro}.[1] Therefore, they were introduced to the LBB for the identification of 3C^{pro}. More importantly, a desired functional group used for the screening reaction should be introduced into the LBB. To avoid the formation of non-biocatalyzed adducts, an ideal screening reaction applied in the method should have a low yield and slow kinetics in the absence of a template protein. Our recent report provided evidence that under biochemical conditions, direct aldol condensation reaction barely proceeds in aqueous buffers.[7] Therefore, aldol condensation reaction was selected as the screening reaction. As mentioned above, LBB (**8**) was designed and synthesized with a terminal aldehyde group as the functional group to screen the P3 fragment of the EV-A71 3C^{pro} inhibitor. (The detailed procedure is given in Part A of the SI).

3C^{pro} was subjected to LBB (**8**) in 1.5 mL plastic tubes with a final protein concentration of 40 μM and an LBB (**8**) concentration of 2.4 mM in PB buffer (25 mM, pH = 7.0). The solution was incubated at room temperature for 12 hours. Subsequently, the small molecules were removed by three ultra-filtration steps to prepare the protease solution for analysis using LC-ESI-Orbitrap MS analysis. As expected, the desired signal of 3C^{pro} modified with **8** was observed, and a

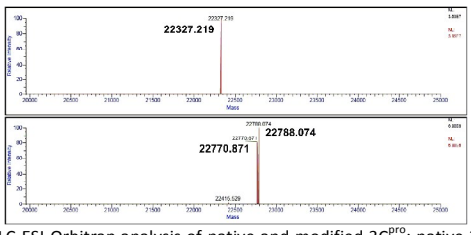


Figure 3. LC-ESI-Orbitrap analysis of native and modified 3C^{pro}: native 3C^{pro} (top) and 3C^{pro} modified with aldehyde **8** (bottom).

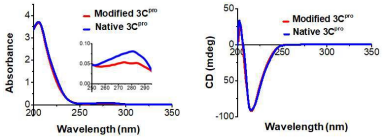


Figure 4. UV/vis (left) and CD (right) spectra of native 3C^{pro} and 3C^{pro} modified with LBB **8.**

conversion rate of 100.0% was quantitatively determined by MS analysis (Figure 3 and Figure S25-26, the detailed procedure is provided in Part D and Part F of the SI).[8] Site-specific modification was further confirmed by trypsin digestion. LC–MS/MS analysis of the resulting peptide fragments revealed that the Michael addition occurred site-specifically with Cys147 at the catalytic site (see Figure S2-S5 in the SI). The comparison of circular dichroism (CD) and UV/vis spectra of the native protease and the protease modified by Michael addition with **8** showed few changes, which provided evidence that the structure/conformation of the protein was not changed after the modification (Figure 4).

To verify the feasibility of the method, 28 commercially available ketones (Figure 5) were screened to investigate whether the aldol reaction occurred in the S3 pocket of 3C^{pro} modified with LBB (**8**). At the beginning, the mixture of several ketones was applied as a group for screening. In such experiments, a small amount of protein deposition was, presumably because the overall concentration of ketones was too high. After the screening of mixed ketones, we screened the 28 ketones indvidually, in order to avoid affecting the protein structure and to enhance the fidelity of the screening results.

The modified protease solution was subjected to different ketones in 1.5 mL plastic tubes with a final protein concentration of 40 μM and a ketone concentration of 80 mM in PB buffer (25 mM, pH= 7.0). The solution was then incubated at room temperature for 12 h. Subsequently, the small molecules and precipitate were removed using three ultra-filtration steps to prepare the protease solution for analysis. Interestingly, the desired signal of modified 3C^{pro} reacted with 4-acetylpyridine was detected. A conversion rate of 40.5% was achieved as quantitatively analyzed by LC-ESI-Orbitrap MS analysis，according to the published protocol (Figure 6 and Figure S27, the detailed procedure is given in Part D and Part F in the SI).[8] In addition, the desired signal of modified 3C^{pro} reacted with 3-acetylpyridine was also detected and the conversion rate was 29.8% (Figure 6 and Figure S28, the detailed procedure is given in Part D and Part F in the SI).[8] Except for 4-acetylpyridine and 3-acetylpyridine, the aldol

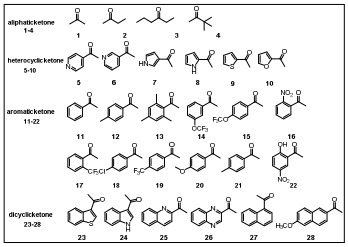


Figure 5. The collection of commercial ketones used for screening.

View Article Online
DOI: 10.1039/C8CC00469B

reaction did not proceed with other tested ketones under the same conditions. Native 3C^{pro} was employed as a control, and no product was observed (see Figure S20 and Figure S21 in the SI).

The site-specific modification of the protease product was further confirmed by trypsin digestion. LC−MS/MS analysis of the resulting peptide fragments showed that the aldol reaction occurred site-specifically between the terminal aldehyde of LBB and 4-acetylpyridine or 3-acetylpyridine (see Figure S6-9 in the SI).

In the absence of EV-A71 3C^{pro}, the reaction of 4-acetylpyridine or 3-acetylpyridine with LBB (**8**) was also performed. LBB (**8**) and 4-acetylpyridine or 3-acetylpyridine were mixed and shaken in the absence of EV-A71 3C^{pro}. The reaction conditions, including the pH, the concentrations and the temperature, were identical with those used for the screening reaction. The aldol reaction, monitored by ^{1}H-NMR and HPLC, did not proceed and no new compound was detected. These data confirmed that the aldol reaction between LLB (**8**) and 4-acetylpyridine or 3-acetylpyridine was driven by EV-A71 3C^{pro} in the desired active site (see Figure S10-19 in SI).

Because the aldol reaction of 4-acetylpyridine or 3-acetylpyridine and LLB (**8**) occurred only in the presence of the 3C^{pro}, 4-acetylpyridine and 3-acetylpyridine were postulated to have higher binding affinity to the protease than other ketones. To verify our assumption, **12a-g**, corresponding to pinacolone, 2'-acetonaphthone, 4-acetylpyridine, 2-acetylthiophene, acetophenone, 3-acetylpyridine, and 1-(2-nitrophenyl)-ethanone were synthesized and evaluated for their enzyme-inhibition activity (Part A of the SI and Table 1).

For the evaluation of inhibition activities of **12a-g**, rupintrivir, with the α,β-unsaturated ester, was used as the positive control. Compounds **12a-g** showed significantly better activeties than **8** (IC_{50} = 16.38 ± 0.12 μM), which suggests that the P3 fragment is important for the inhibitors. As shown in Table 1, **12c** (IC_{50} = 1.43 ± 0.10 μM) presented approximately 4- to 5-fold better activity than **12a** (IC_{50} = 6.23 ± 0.20 μM), **12b** (IC_{50} = 6.54 ± 0.21 μM), **12d** (IC_{50} = 6.20 ± 0.13 μM), **12e** (IC_{50} = 6.39 ± 0.13 μM), and **12g** (IC_{50} = 6.72 ± 0.12 μM). In addition, compared with **12c**, **12f** (IC_{50} = 1.61 ± 0.07 μM) showed lower activity**,** but it presented approximately 4-fold better activity than **12a**, **12b**, **12d**, **12e** and **12g**. Furthermore, the first-order rate constant (k_{obs}) was introduced to verify the comparison of the obtained enzyme-inhibition activity (IC_{50}). As listed in Table 1, the k_{obs}/[I] (**12a**-**g**) showed good correlation with IC_{50}. These results were consistent with the screening results, supporting

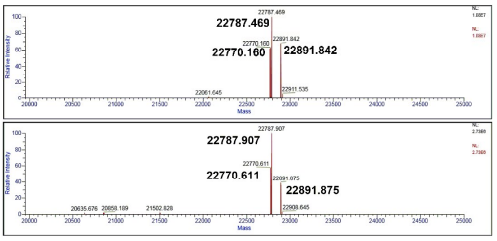
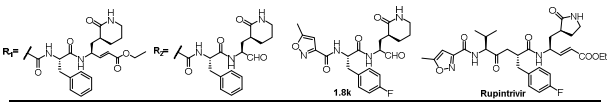
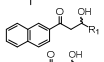
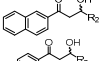


Figure 6. LC-ESI-Orbitrap analysis of the aldol reaction of modified protease with 4-acetylpyridine (top); LC-ESI-Orbitrap analysis of the aldol reaction of modified protease with 3-acetylpyridine (bottom).

Table1. Enzyme-inhibition activity of **8**，**12a-12g** and **17a-17g** against EV-A71 3C^{pro}



NO		IC_{50}[a, b] (μM)	k_{obs}/[I] (M^{-1} s^{-1})[c]
8		16.38 ± 0.12	5347 ± 121
12a		6.23 ± 0.20	10320 ± 372
12b		6.54 ± 0.21	12743 ± 425
12c		1.43 ± 0.10	216347± 5752
12c-1		1.41 ± 0.05	236307 ± 7230
12c-2		1.44 ± 0.07	182500 ± 5607
12d		6.20 ± 0.13	21586 ± 794
12e		6.39 ± 0.13	14680 ± 472
(7S,2'S,10S, 13R)-12e		6.42 ± 0.15	12549 ± 289
(7S,2'S,10S, 13S)-12e		6.34 ± 0.13	16388 ± 476
12f		1.61 ± 0.07	103887 ± 3346
12g		6.72 ± 0.12	21327 ± 689
17a		1.52 ± 0.02	
17b		1.41 ± 0.01	
17c		0.15 ± 0.01	
17d		1.45 ± 0.03	
17e		1.37 ± 0.02	
(2S,2'S,5S,8 R)-17e		1.36 ± 0.02	
(2S,2'S,5S,8 S)-17e		1.47 ± 0.01	
17f		0.17 ± 0.01	
17g		1.55 ± 0.03	
Rupintrivir[d]	-----	1.13 ± 0.02	615418 ± 22517
1.8k[e]	-----	0.32 ± 0.02	

[a]The pH of the biochemical assay for the inhibitors was 7.0. [b]All values are the mean of three independent experiments. [c]Analytical methods were carried out as reported in the literature.[9] [d]Rupintrivir was used as the positive control.[e]Compound **1.8k**, as reported in the reference, was used as the positive control.[1]

the idea that 4-acetylpyridine or 3-acetylpyridine are better fit with the S3 pocket than the other commercial ketones. Such results also support the feasibility of TGSOF in screening for a fragment of the desired inhibitor. To elucidate the binding mode of **12c** and **12f** with EV-A71 3C^{pro}, a docking study was conducted using MOE (see Figure S22 in the SI). The striking difference to other ketones was that the 4-acetylpyridine of **12c** or 3-acetylpyridine of **12f** interacted with Asn165. Because the nitrogen atom of 4-acetyl pyridine or 3-acetylpyridine is electron-rich, it formed a hydrogen bond with Asn165. For **12a**, **12b**, **12d**, **12e** and **12g** such an interaction was not observed. This specific binding element might be critical for the activity difference, resulting in **12c** or **12f** exhibiting better activity. Further, compared with **12c**, the hydrogen bond distance

View Article Online
DOI: 10.1039/C8CC00469B

between **12f** with Asn165 was a little farther away (2.2 Å for **12f** vs. 1.9 Å for **12c**), so **12f** exhibited slightly lower activity.

To define the absolute configuration of the two isomers of the aldol reaction product, the epimeric mixture **12e** was separated to yield **(7S,2'S,10S,13R)-12e** and **(7S,2'S,10S,13S)-12e**. The absolute configuration of the chiral carbon of the hydroxyl group was determined by the modified Mosher's method.[10] The result was further supported by ^{1}H-NMR (see Part B in SI). The epimeric mixture **12e** (IC_{50} = 6.39 ± 0.13 μM) and the isolated isomers **(7S,2'S,10S,13R)-12e** (IC_{50} = 6.42 ± 0.15 μM) and **(7S,2'S,10S,13S)-12e** (IC_{50} = 6.34 ± 0.13 μM) exhibited equipotent inhibitory activity against EV-A71 $3C^{pro}$. Similar results were observed for the epimeric mixture **12c** (IC_{50} = 1.43 ± 0.10 μM) and both isomers **12c-1** (IC_{50} = 1.41 ± 0.05 μM) and **12c-2** (IC_{50} = 1.44 ± 0.07 μM).

The interactions of epimeric mixture **12c** docked to EV-A71 $3C^{pro}$ could explain the unsubstantial difference in the IC_{50} values. As shown in Figure S24, the two isomers had the same interaction distances between the 4-acetylpyridine nitrogen atom and Asn165 (1.9 Å). Moreover, the OH oxygen of the two isomers could form a similar weak hydrogen bond with Gly164 (2.8 Å vs. 2.9 Å).

To obtain more potent inhibitors, we then optimized the structure of **12c** and **12f**. Based on the characterization of the binding modes, the inhibitors can be divided into reversible inhibitors (e.g., aldehydes) and irreversible inhibitors (e.g., α,β-unsaturated ester).[11] The previous study of EV-A71 $3C^{pro}$ inhibitors provided evidence that the aldehyde group fits readily into the S1'pocket of EV-A71 $3C^{pro}$ and the interaction is more favorable than the unsaturated ester. Therefore, the warhead α,β-unsaturated ester of **12c** and **12f** were replaced by an aldehyde group to obtain **17c** and **17f**. In addition, **17a**, **17b**, **17d, 17e** and **17g** corresponding to pinacolone, 2'-acetonaphthone, 2-acetylthiophene, acetophenone and 1-(2-nitrophenyl)-ethanone were also synthesized and evaluated for their enzyme-inhibition activity (Table 1 and Part A of the SI). Similar to examination of the absolute configuration of the two isomers of **12e,** the epimeric mixture **17e** was separated to yield **(2S,2'S,5S,8R)-17e** and **(2S,2'S,5S,8S)-17e** (Part B in SI).

For reversible inhibitors **17a-g**, **1,8k** with the aldehyde group reported in previous literature, was applied as the positive control. The enzyme-inhibition activity indicated that aldehyde **17c** (IC_{50} = 0.15 ± 0.01 μM) exhibited approximately 10-fold more potent activity than **12c,** and **17f** (IC_{50} = 0.17 ± 0.01 μM) exhibited similar result in the protease assay. Compared to the α,β-unsaturated ester, the aldehyde group formed a reversible covalent bond with Cys147, which fit in the catalytic site of EV-A71 $3C^{pro}$. In addition, as shown in Table 1, **17c** and **17f** presented approximately 9-fold better activity than **17a** (IC_{50} = 1.52 ± 0.02 μM), **17b** (IC_{50} = 1.41 ± 0.01 μM), **17d** (IC_{50} = 1.45 ± 0.03 μM), **17e** (IC_{50} = 1.37 ± 0.02 μM), and **17g** (IC_{50} = 1.55 ± 0.03 μM). These findings verified that 4-acetylpyridine and 3-acetylpyridine were still the more potent P3 fragments for the inhibitors with aldehyde warhead. Therefore, based on **12c** and **12f**, we successfully obtained **17c** and **17f** as reversible inhibitors with excellent enzyme-inhibition activity against EV-A71 $3C^{pro}$.

In summary, target-guided screening of fragments (TGSOF) has been employed in the study of EV-A71 $3C^{pro}$ inhibitors. We identified 4-acetylpyridine and 3-acetylpyridine as effective P3 fragments and obtained the corresponding irreversible inhibitor **12c** (IC_{50} = 1.43 ± 0.10 μM) and **12f** (IC_{50} = 1.61 ± 0.07 μM) using the method. In view of our previous study of EV-A71 $3C^{pro}$ inhibitors and based on **12c** and **12f,** the reversible inhibitor **17c** and **17f** were obtained, and they exhibited excellent inhibitory activity (IC_{50} = 0.15 ± 0.01 μM and IC_{50} = 0.17 ± 0.01 μM). These findings strongly support the use of TGSOF as an efficient way to screen for a specific inhibitor fragment and as a novel strategy to discover lead compounds.

In TGSOF, the functional group used to screen fragment can be introduced at a specific site of the target biomolecule using LBB. Furthermore, the screening reaction of the modified protein and fragments is to happen in a specific location of interest. TGSOF can be employed to screen other fragments of EV-A71 $3C^{pro}$ inhibitors, and then novel inhibitors could be obtained by the assembly of those screened fragments. Therefore, TGSOF offer a way to rapidly identify effective fragments for drug discovery. For future studies, TGSOF might be applied to study other cysteine proteases, such as cathepsins. Moreover, instead of aldol condensation reaction, other reactions can also be applied in TGSOF (i.e. the substitution reaction of C–N bond formation, the substitution reaction of C–S bond formation).[4]

This work was supported by the National Basic Research Program of China (973 program, Grant Nos. 2013CB911104, 2013CB911100), the National Natural Science Foundation of China (Grant Nos. 21672115), and the "111" Project of the Ministry of Education of China (Project No. B06005).

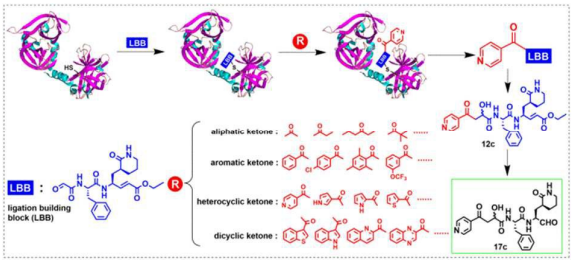
Notes and references

1 Y. Y. Zhai, Y. Y. Ma, F. Ma, Q. D. Nie, X. J. Ren, Y. X. Wang, L. Q. Shang and Z. Yin, *Eur. J. Med. Chem.*, 2016, **124**, 559-573.
2 L. Q. Shang, M. Y. Xu and Z. Yin, *Antiviral Res.*, 2013, **97**, 183-194.
3 Y. Y. Zhai, X. S. Zhao, Z. J. Cui, M. Wang, Y. X. Wang, L. F. Li, Q. Sun, X. Yang, D. B. Zeng, Y. Liu, Y. N. Sun, Z. Y. Lou, L. Q. Shang and Z. Yin, *J. Med. Chem.*, 2015, **58**, 9414-9420.
4 X. D. Hu and R. Manetsch, *Chem. Soc. Rev.*, 2010, **39**, 1316-1324.
5 (a) Y. Wang, J. J. Zhu and L. W. Zhang, *J. Med. Chem.*, 2017, **60**, 263-272; (b) E. Oueis, C. Sabot and P. Y. Renard, *Chem. Commun.*, 2015, **51**, 12158-12169; (c) D. Becker, Z. Kaczmarska, C. Arkona, R. Schulz, C. Tauber, G. Wolber, R. Hilgenfeld, M. Coll and J. Rademann, *Nat Commun*, 2016, **7**.
6 A. Herrmann, *Chem. Soc. Rev.*, 2014, **43**, 1899-1933.
7 P. Wang, S. M. Zhang, Q. Y. Meng, Y. Liu, L. Q. Shang and Z. Yin, *Org. Lett.*, 2015, **17**, 1361-1364.
8 Q. Ruan, Q. C. Ji, M. E. Arnold, W. G. Humphreys and M. S. Zhu, *Anal. Chem.*, 2011, **83**, 8937-8944.
9 J. Z. Tan, S. George, Y. Kusov, M. Perbandt, S. Anemuller, J. R. Mesters, H. Norder, B. Coutard, C. Lacroix, P. Leyssen, J. Neyts and R. Hilgenfeld, *J. Virol.*, 2013, **87**, 4339-4351.
10 F. Freire, J. M. Seco, E. Quinoa and R. Riguera, *J. Org. Chem.*, 2005, **70**, 3778-3790.
11 H.-H. Otto and T. Schirmeister, *Chem. Rev.*, 1997, **97**, 133-172.

View Article Online
DOI: 10.1039/C8CC00469B

View Article Online
DOI: 10.1039/C8CC00469B

Graphical Abstract



Target-guided screening of fragments (TGSOF) was developed and employed in the identification of EV-A71 3C protease ($3C^{pro}$) inhibitors.