



Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Virtual screening and further development of novel ALK inhibitors

Masako Okamoto^{a,b}, Hirotatsu Kojima^b, Nae Saito^b, Takayoshi Okabe^b, Yoshiaki Masuda^{a,b}, Toshio Furuya^{a,b}, Tetsuo Nagano^{b,c,*}

^a Drug Discovery Department, Research & Development Division, PharmaDesign, Inc., 2-19-8 Hatchobori, Chuo-ku, Tokyo 104-0032, Japan

^b Chemical Biology Research Initiative, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

^c Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

ARTICLE INFO

Article history:

Received 3 March 2011

Revised 2 April 2011

Accepted 4 April 2011

Available online 7 April 2011

Keywords:

ALK, anaplastic lymphoma kinase
NSCLC, non-small-cell lung cancer
Virtual screening
Chemical library

ABSTRACT

Anaplastic lymphoma kinase (ALK) has been in the spotlight in recent years as a promising new target for therapy of non-small-cell lung cancer (NSCLC). Since the identification of the echinoderm microtubule-associated protein-like 4 (EML4)-ALK fusion gene in some NSCLC patients was reported in 2007, various research groups have been seeking ALK inhibitors. Above all, crizotinib (PF-02341066) has been under clinical trial, and its therapeutic efficacy of inhibiting ALK in NSCLC has been reported. Among anticancer drugs, drug resistance appears frequently necessitating various kinds of inhibitors. We identified novel ALK inhibitors by virtual screening from the public chemical library collected by the Chemical Biology Research Initiative (CBRI) at the University of Tokyo, and inhibitors that are more potent were developed.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase and is normally expressed in the central and peripheral nervous systems. ALK was first identified as a fusion partner of nucleophosmin (NPM) in anaplastic large-cell lymphomas (ALCLs) with a t(2;5)(p23;q35) chromosomal rearrangement.¹ Hence, ALK has been expected to be a target for anticancer therapy similar to other tyrosine kinases. Various research of ALK inhibitors has been conducted to date; for example, a small molecule ALK inhibitor, NVP-TAE684, was reported by The Genomics Institute of the Novartis Research Foundation, showing potent and selective inhibitory activity.²

Moreover, in 2007, the transforming echinoderm microtubule-associated protein-like 4 (EML4)-ALK fusion gene in non-small-cell lung cancer (NSCLC) was reported by Mano et al.³ ALK has since been in the spotlight as a promising new target for therapy of non-small-cell lung cancer (NSCLC). Since this report, many research groups have been seeking ALK inhibitors and have reported some potent inhibitors. Above all, crizotinib (PF-02341066)⁴ produced by Pfizer, Inc., has already been under clinical trial; its IC₅₀ value is 24 nM against NPM-ALK oncogenic fusion variant of the ALK receptor tyrosine kinase expressed by the KARPAS299 human

anaplastic large cell lymphoma (ALCL) cell line, and its therapeutic efficacy of inhibiting ALK and c-Met in NSCLC is truly remarkable. However, since drug resistance appears frequently for anticancer drugs, various kinds of inhibitors are still required.

To obtain novel ALK inhibitors having new scaffolds, we adopted an *in silico* approach. In the past several years, we have attempted to find various biologically active compounds against each target protein from a commercially available compound database using *in silico* approach and have reported successful results.^{5,6} Furthermore, a public chemical library has been collected and supplied by the Chemical Biology Research Initiative (CBRI)⁷ at the University of Tokyo in Japan since 2007; we also used this chemical library and virtual screening approach.

Notably, we have had many successful experiments finding hit compounds using structure-based virtual screening (SBVS). We usually construct protein–ligand complex models using an X-ray crystal structure of the target protein and known inhibitors before SBVS for an efficient screen. However, at the time of starting this research, there were no known X-ray crystal structures of the ALK kinase domain. Therefore, we obtained a 3D structure of the ALK kinase domain by homology modeling using one of the crystal structures of the activated insulin receptor tyrosine kinase (PDB code: 1IR3⁸) as a template. Protein–ligand complex models were then made using this homology model protein structure and known inhibitors, and docking calculations were conducted using CONSENSUS-DOCK.⁹

As a result, we were successful in the identification of hit compounds as novel ALK inhibitors by SBVS from the public chemical library collected by CBRI at the University of Tokyo. In addition,

* Corresponding author. Tel.: +81 3 5841 4850; fax: +81 3 5841 4855.

E-mail addresses: okamoto@pharmadesign.co.jp (M. Okamoto), kojima@mol.f.u-tokyo.ac.jp (H. Kojima), nae-saito@mol.f.u-tokyo.ac.jp (N. Saito), tokabe@mol.f.u-tokyo.ac.jp (T. Okabe), masuda@pharmadesign.co.jp (Y. Masuda), furuya@pharmadesign.co.jp (T. Furuya), tlong@mol.f.u-tokyo.ac.jp (T. Nagano).

we succeeded in the organic synthesis of lead compounds from hit compounds to obtain more potent inhibitors.

2. Results and discussion

2.1. Sequence alignment and homology modeling

Protein 3D structures are indispensable for SBVS; however, X-ray crystal structures had not been included in the PDB database¹⁰ until the release of three crystal structures analyzed by several groups.¹¹ Therefore, we obtained the 3D structure of ALK kinase domain by homology modeling using one of the crystal

structures of activated insulin receptor tyrosine kinase as a template.

To select the template for homology modeling, we practiced BLAST¹² search against PDB database using ALK kinase domain sequence (a part of Q9UM73 (Genbank Accession code)). The result from BLAST search is shown in Table 1, and some crystal structures of IGF-1R and INSR were high-ranking. In particular, INSR (PDB code: 1IR3), a ligand-complex crystal structure having high resolution (1.90 Å), had been used as a template for homology modeling in other papers.^{13,14} Therefore we selected 1IR3 as our template.

Sequence alignment and homology modeling were carried out by MOE¹⁵ module. The alignment is shown in Figure 1. In general, gatekeeper residue is considered important in the role of obtaining selectivity. In the case of this homology model of ALK, Leu (red letter in Fig. 1) was predicted as the gatekeeper, as found by Gunby et al.¹³ After the alignment, we constructed a modeling structure with 1IR3 as a template using the MOE module.

Table 1
BLAST search results using ALK kinase domain sequence as a query

PDB code	Release	Protein	Ligand	%ID
1M7N	2003/1/7	IGF-1R	—	46
1P4O	2003/4/29	IGF-1R	—	45
1JQH	2002/4/19	IGF-1R	ANP	45
2OJ9	2007/5/1	IGF-1R	BMI	45
3D94	2008/7/29	IGF-1R	D94	45
1IRK	1995/2/27	INSR	—	45
1P14	2003/7/22	INSR	—	45
1K3A	2001/11/28	IGF-1R	ACP	44
2Z8C	2008/8/12	INSR	S91	44
1IR3	1998/1/7	INSR	ANP	44
1RQQ	2003/12/30	INSR	112	44
2ZM3	2008/6/10	IGF-1R	575	43
3BKB	2007/12/25	v-Fes	STU	42
3CD3	2008/3/25	v-Fes	STU	42
1I44	2001/3/7	INSR	ACP	41
1LUF	2002/9/11	MuSK	—	40
3LCK	1997/12/3	LCK	—	40
3C4F	2008/2/26	FGFR2	C4F	39
1IEP	2001/4/18	c-Abl	STI	39
1OPJ	2003/4/8	c-Abl	STI	39

2.2. Protein–ligand complex model

For more efficient SBVS, we usually construct a protein–ligand complex model before virtual screening.⁵ We use an X-ray crystal structure ordinarily, but in this case, because crystal structures were lacking at the time of conducting this research, we adopted the model structure of ALK kinase domain described above. In the constructed complex model, we used known potent ALK inhibitors selected from patent information¹⁶ (Table 2). Various ALK inhibitors such as 3H-pyrazolo[3,4-c]isoquinoline derivatives were described abundantly in international publication patent information,¹⁶ and we were able to consider structure–activity relationships from these compounds' assay data. Therefore we selected compounds in Table 2 to build protein–ligand complex models. Then, we tried to construct appropriate protein–ligand complex models for efficient virtual screening.

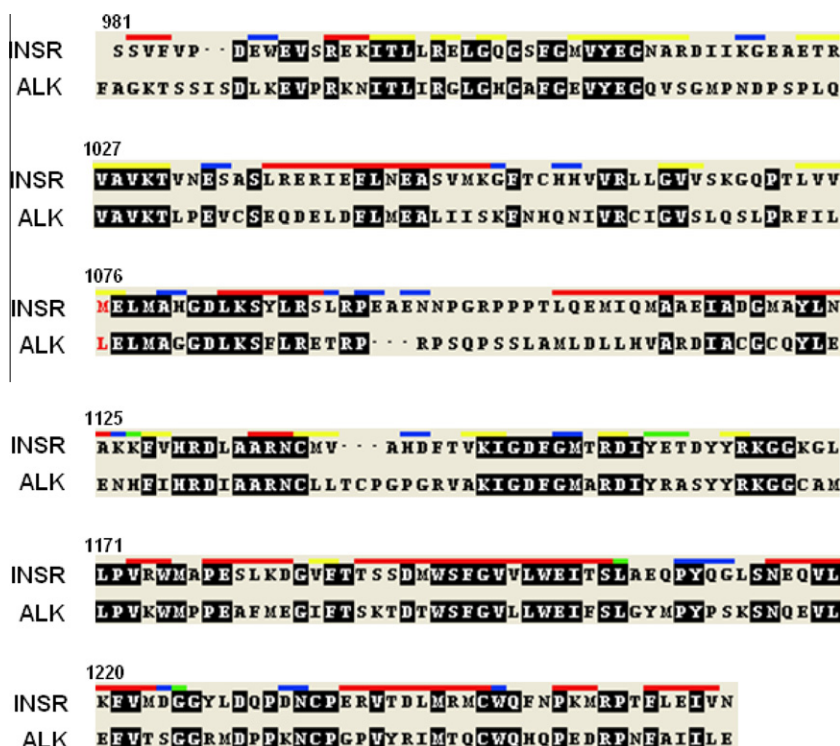
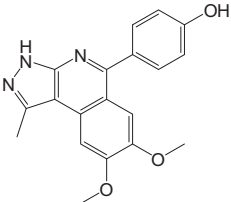
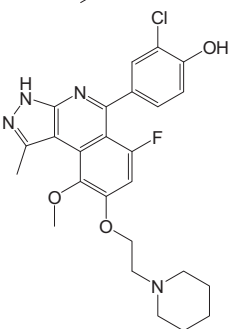


Figure 1. Sequence alignment of INSR and ALK kinase domain. Numbers are residue UID of INSR. Color bars show actual secondary structure of INSR (PDB code: 1IR3). Reversed letters indicate conserved residues. The gatekeeper residues are shown in red letters. This figure was prepared by MOE.

Table 2
Known potent ALK inhibitors¹⁶ used for building protein–ligand complex models

Compound	Structure	Inhibitory activity
1		IC ₅₀ ≤ 99 nM
2		IC ₅₀ ≤ 99 nM

In this study, we used a molecular dynamics simulation module in MOE¹⁵ by modifying an SVL script to build the complex models. This SVL script, called MultiCopyMD,⁵ was devised to enable several ligand molecules to be positioned in the binding site of the target protein simultaneously during the simulation so that the consensus binding conformation of that protein for multiple ligands can be generated. Water molecules were soaked within 20 Å of the center of the ligands in the complex model with wall restraints around them, while the side chain atoms within 4.5 Å of the center of the ligands and the backbone atoms of the glycine-rich loop were unfixed in the MD calculations. The system was gradually heated to 300 K, and an additional 1 ns simulation at constant temperature and volume (NVT ensemble, NPA algorithm) was carried out. We then clustered the coordinates in this trajectory and selected several structures to compare each enrichment of virtual screening.

We compared the enrichments among representative complex model structures by small-scale (approximately a 1000-compound test set database) virtual screening using CONSENSUS-DOCK.⁹ These 1000 compounds were selected randomly from the commercially available compound databases, filtered by drug-likeness and clustering. The small-scale test set database had these 1000 compounds and known ALK inhibitors, and a complex model structure having the best enrichment of known ALK inhibitors in small-scale virtual screening was selected.

The most appropriate complex model structure selected is shown in Figure 2. In this binding mode, there are four hydrogen bonds between compound 2 and ALK kinase domain including two hydrogen bonds with hinge region, which seem to be core interactions for most kinase inhibitors; hydrophobic interaction is also predicted with Leu1196.

2.3. Structure-based virtual screening

Using the complex model described above, we conducted structure-based virtual screening using CONSENSUS-DOCK⁹ against the CBRI Library and commercially available compound database. At the time we carried out these screenings, the CBRI Library contained 71,558 compounds. To compensate for the numbers of compounds for structure-based virtual screening, we also screened against a commercially available compound database. After docking calculation, we selected compounds by pharmacophore that

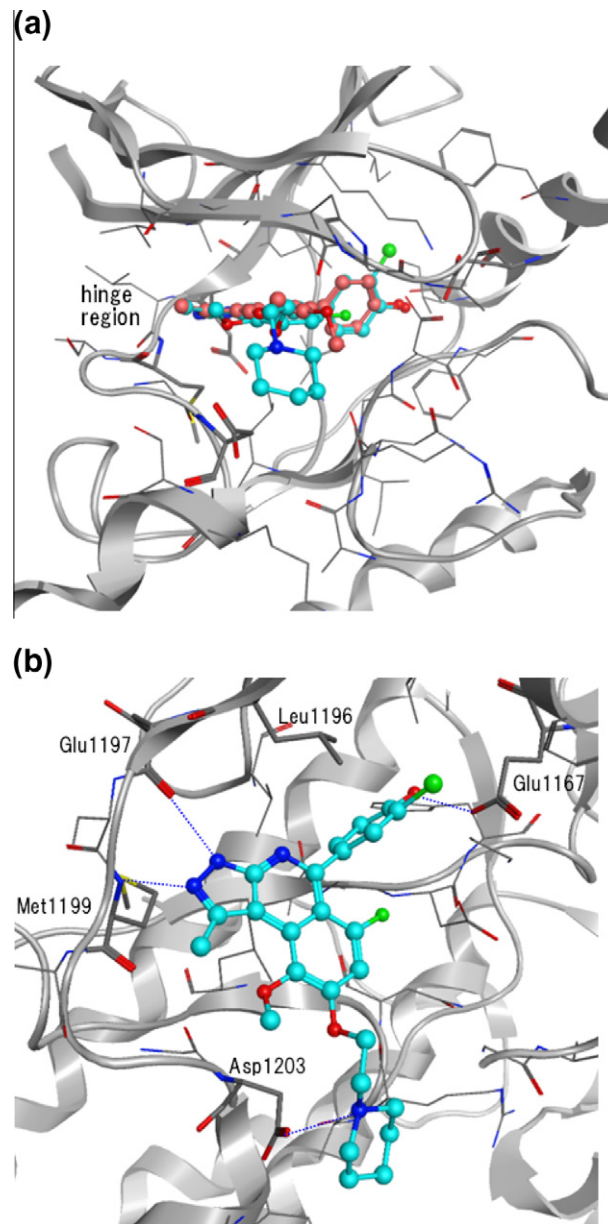


Figure 2. The protein–ligand complex model structures of ALK kinase domain. (A) One known inhibitor, compound 1, is shown in pink; and the other known inhibitor, compound 2, is shown in light blue. (B) To show detailed interactions, only compound 2 is drawn. There are four hydrogen bonds between compound 2 and ALK kinase domain, and a hydrophobic interaction is also predicted with Leu1196.

could possibly make hydrogen bonds with the hinge region of the ATP binding site.

Through SBVS, as shown in Figure 3, we selected 300 compounds from CBRI Library and 591 compounds from the commercially available compound database. Among these compounds, we performed an ALK inhibition assay. As a result, 13 compounds in total exhibited greater than 50% inhibition at 10 μM, as shown in Table 3. To obtain more potent compounds, we conducted a similarity search using these compounds as queries.

2.4. Second screening (similarity search)

Using the 13 hit compounds as queries, we carried out a similarity search with BIT_MACCS fingerprint and selected 6036 compounds. As a result, we were able to obtain 64 compounds with equivalent or higher inhibitory activity (% inhibition at

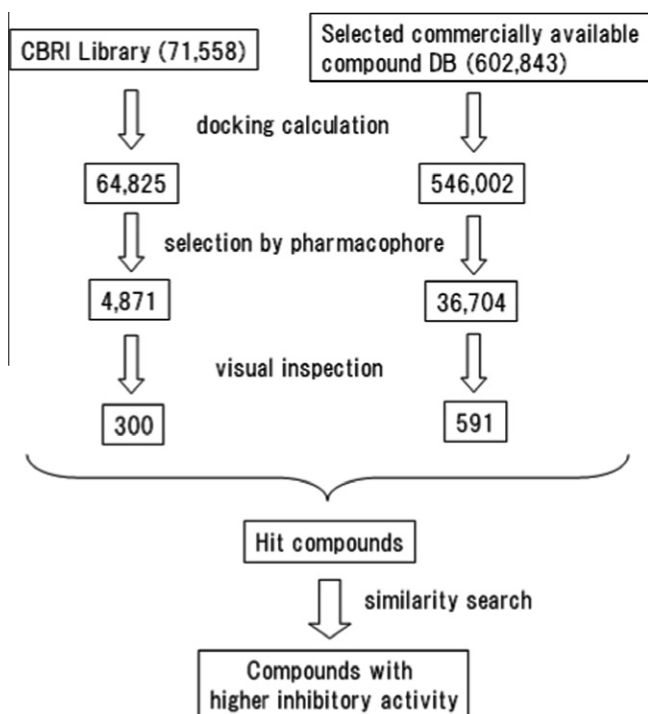


Figure 3. Structure-based virtual screening protocol.

10 μ M > 50). IC_{50} value was determined to find relatively potent compounds. Fifteen compounds showed an IC_{50} value under 10 μ M (Table 4); the most potent compound's IC_{50} value was 0.85 μ M (compound 3).

The predicted binding mode of compound 3 is shown in Figure 4. In this binding mode, there were two hydrogen bonds between compound 3 and the hinge region of ALK kinase domain; these interactions were thought to stabilize the binding of compound 3. Intramolecular hydrogen bonding was also predicted between

the O of indolinone and NH of pyrrole. Furthermore, hydrophobic interactions between compound 3 and Leu1196, Leu1256 were predicted. Based on these points, indolinone was predicted as a promising scaffold for further development.

2.5. Design and synthesis

For further development, we planned to synthesize designed compounds. To add more interactions between compound 3 and adjacent amino acids of the ATP binding site outside of the hinge region, we synthesized compounds in Table 5 as shown in Schemes 1–3; detail protocols are described in Section 4.

We conducted in vitro assays of the synthesized compounds, and the results are shown in Table 6. The most potent inhibitor was compound 25a; its predicted binding mode is shown in Figure 5. Compared to the binding mode of compound 3, two additional hydrogen bonds were predicted: NH_2 on the indolinone ring and O in the side chain of Asp178, and NH_2 on the benzene ring and O in the side chain of Asp111. These interactions were thought to increase inhibitory activity of compound 25a. According to the comparison between 25a and 25b, *m*-position of amino group on benzene ring seemed to be better than *p*-position. We designed compounds 26a and 26b to obtain additional hydrogen bonds, but those inhibitory activities decreased. This is due to the exposure of the (2-aminoethyl)amino group to solvent, but hydrogen bonds were not made with neighboring amino acids.

2.6. Kinase panel assay

To examine kinase selectivity of synthesized compound 23a, we conducted kinase panel assay including 21 tyrosine kinases (Table 7). As a result, 23a showed highly selectivity against ABL, CSK, EGFR, EPHA2, EPHB4, and JAK3. However, against FLT3, IGF-1R, INSR, TRKA, and TYRO3, inhibitory activities of 23a were equal to ALK, EML4-ALK, and NPM1-ALK. Especially, selectivity against INSR is essential for drugs as ALK inhibitors, we have been considering drug design for obtaining selective compounds.

Table 3

Hit compounds from structure-based virtual screening (% inhibition at 10 μ M > 50)

1st hit compounds from SBVS				

Table 4Compounds with high inhibitory activity in 2nd screening ($IC_{50} \leq 10 \mu M$)

Compound	Structure	IC_{50} (μM)	Compound	Structure	IC_{50} (μM)
3		0.85	10		2.85
4		0.91	11		3.47
5		1.38	12		3.75
6		1.73	13		4.24
7		1.81	14		6.46
8		2.40	15		8.66
9		2.40	16		9.29

3. Conclusion

Using SBVS, we identified novel ALK inhibitors from the CBRI Library and a commercially available compound database; these hit compounds had new scaffolds for ALK inhibitor development. Furthermore, we developed lead compounds from hits by designing compounds, and we were successful in obtaining potent ALK inhibitors (the most potent inhibitor: compound **25a**, IC_{50} = 50.8 nM).

The results described in this paper indicate that our SBVS approach is very efficient for finding a novel scaffold on a target protein's active compounds. After finding hit compounds, we could confirm improvements in inhibitory activity through stepwise structure-based drug design. We conducted a kinase panel study to inspect kinase selectivity, and we would like to develop our novel ALK inhibitors. We hope our discovery of these promising ALK inhibitors possessing new scaffolds will be useful for research of chemotherapy for NSCLC.

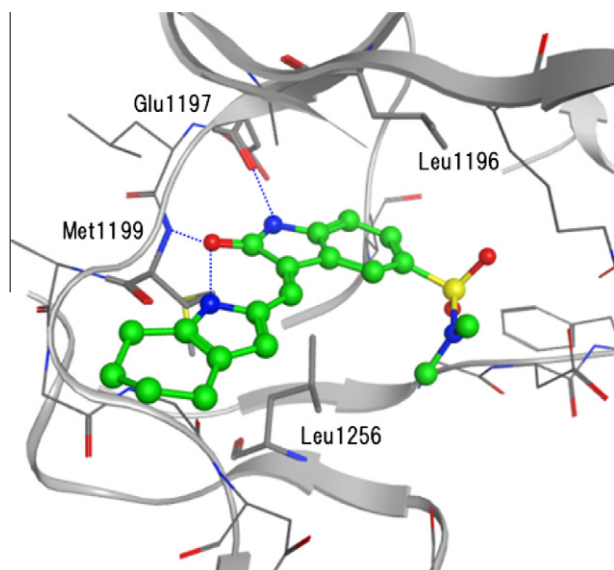


Figure 4. Predicted binding mode of compound 3.

4. Experimental

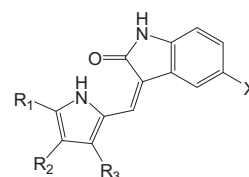
4.1. Compounds

Tested compounds in Table 2 and 3 were purchased from several suppliers. Their characterization and purity were confirmed using LC–MS.

4.2. Similarity search

We calculated a fingerprint based on the descriptors of the BIT_MACCS: MACCS structural keys for all compounds in CBRI

Table 5
Synthesized compounds

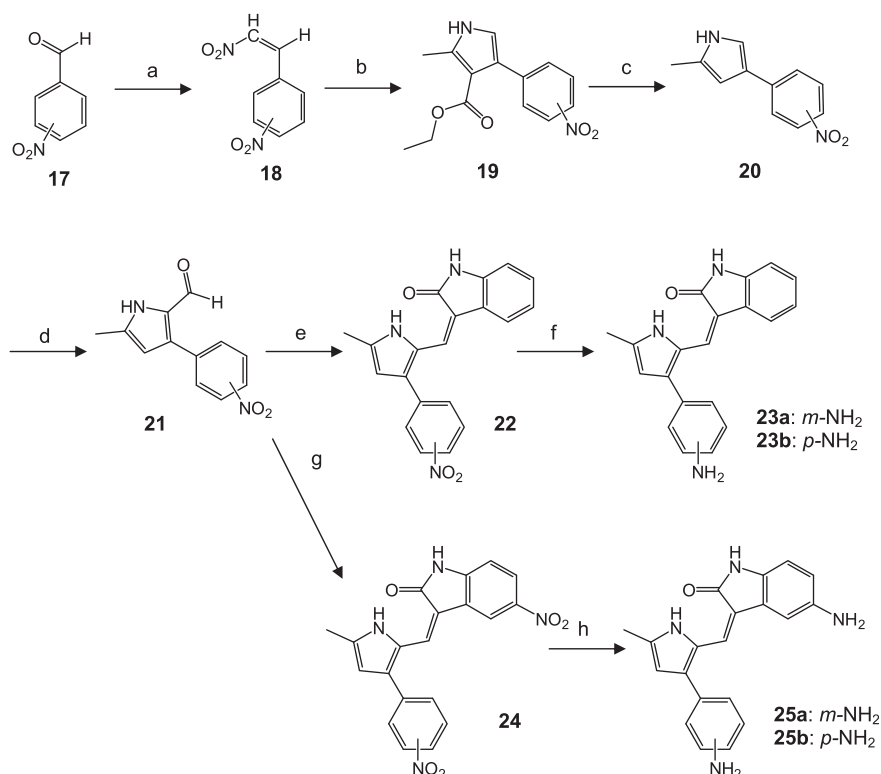


Compound	X	R ₁	R ₂	R ₃
23a	H	CH ₃	H	<i>m</i> -NH ₂ -C ₆ H ₄
23b	H	CH ₃	H	<i>p</i> -NH ₂ -C ₆ H ₄
25a	NH ₂	CH ₃	H	<i>m</i> -NH ₂ -C ₆ H ₄
25b	NH ₂	CH ₃	H	<i>p</i> -NH ₂ -C ₆ H ₄
26a	H	CH ₃	H	<i>m</i> -NH(CH ₂) ₂ NH ₂ -C ₆ H ₄
26b	H	CH ₃	H	<i>p</i> -NH(CH ₂) ₂ NH ₂ -C ₆ H ₄
31a	H	CH ₃	<i>m</i> -NH ₂ -C ₆ H ₄	CH ₃
31b	H	CH ₃	<i>p</i> -NH ₂ -C ₆ H ₄	CH ₃
32a	H	CH ₃	<i>m</i> -NH(CH ₂) ₂ NH ₂ -C ₆ H ₄	CH ₃
32b	H	CH ₃	<i>p</i> -NH(CH ₂) ₂ NH ₂ -C ₆ H ₄	CH ₃

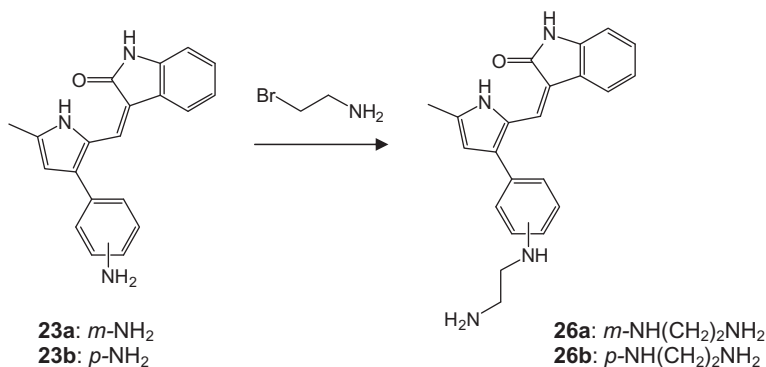
Library; a similarity search was conducted using MOE¹⁵ with compounds in Table 3 as queries. Compounds with similarity were selected by visual inspection for in vitro assay.

4.3. Kinase assay

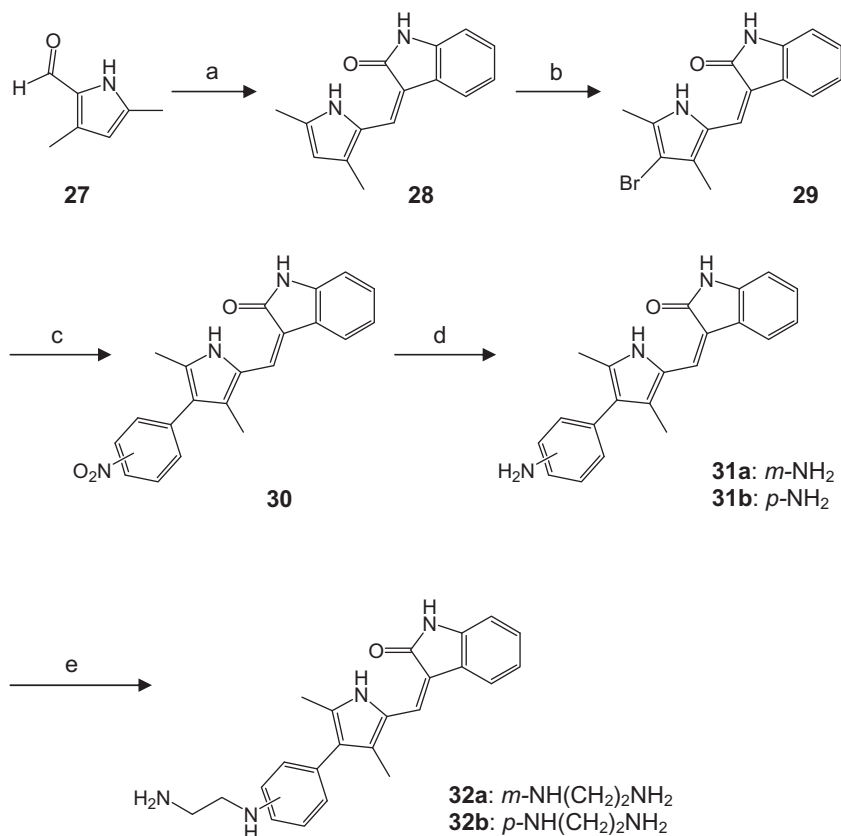
Activity was measured by mobility shift assay. The reaction mixture containing 1.5 μM FL-Peptide 13, 5-FAM-KKSRGDYMTM-QIG-CONH₂ (Caliper Life Science, Hopkinton, MA), 67.9 μM ATP, 42 ng/mL ALK (Invitrogen, Carlsbad, CA), 50 mM Hepes (pH 7.4), 10 mM MgCl₂, 1 mM DTT, 1% Protease Inhibitor Cocktail Set V (Merck Darmstadt, Germany), 1% Phosphatase Inhibitor Cocktail Set III (Merck), 0.01% Brij-35 and a test compound (5% DMSO) was incubated at room temperature for 60 min, and the reaction



Scheme 1. Synthetic protocol of compounds 23a, 23b, 25a, and 25b in Table 5. Reagents: (a) CH₃CO₂NH₄, AcOH, CH₃NO₂; (b) ethyl 3-oxobutanoate, MeOH, CH₃ONa, NH₃/MeOH; (c) TFA; (d) DMF/POCl₃, 1,2-dichloroethane; (e) indoline-2-one, MeOH, piperidine; (f) Pd/C, THF; (g) 5-nitroindoline-2-one, MeOH, piperidine; (h) Pd/C, THF.



Scheme 2. Synthetic protocol of compounds **26a**, **26b** in Table 4. Reagents: 2-bromoethylamine hydrobromide, K₂CO₃, DMF.



Scheme 3. Synthetic protocol of compounds **31a**, **31b**, **32a**, and **32b** in Table 5. Reagents: (a) indoline-2-one, MeOH, piperidine; (b) NBS, (C₆H₅CO)₂O₂, CCl₄; (c) Suzuki coupling, 4-nitrophenylboronic acid, K₂CO₃, Pd(dppf)Cl₂, H₂O, dioxane; (d) Pd/C, THF; (e) 2-bromoethylamine hydrobromide, K₂CO₃, DMF.

was stopped by adding 140 mM EDTA. The phosphorylated and unphosphorylated peptides were separated and detected by Lab-Chip EZ Reader II (Caliper Life Science).

4.4. Kinase panel assay

This experiment was performed by Carna Biosciences (Kobe, Japan). Briefly, the reaction mixture containing the enzyme, 1 mM ATP and 1 μM appropriate peptide was analyzed by mobility shift assay after incubation.

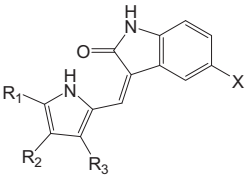
4.5. Chemistry

Unless otherwise noted, materials were obtained from commercial suppliers and used without further purification. ¹H NMR spec-

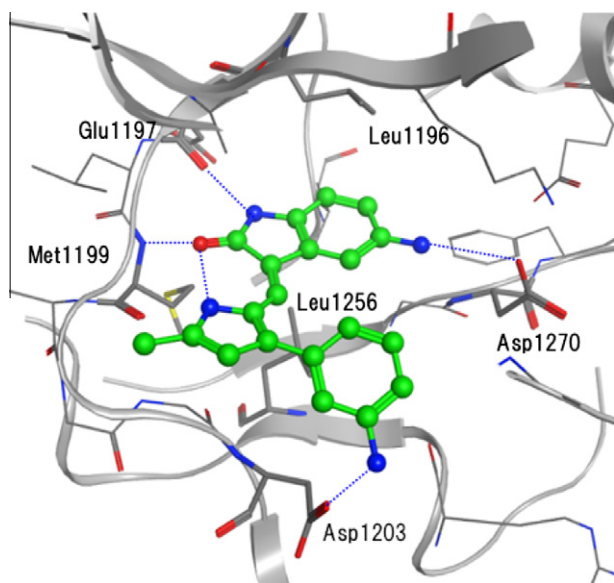
tra were recorded at 300 MHz, and TMS was used as an internal standard. High resolution mass spectra were obtained using JEOL AccuTOF JMS-T100LC and an electrospray ionization (ESI) source. LC-MS was taken on a quadrupole Mass Spectrometer on Agilent LC/MSD 1200 Series (Column: Welchom XB-C18 (50 × 4.6 mm, 5 μm) operating in ES (+) or (−) ionization mode; *T* = 30 °C; flow rate = 1.5 mL/min. Prep-HPLC was performed under the following conditions: Welchom C18 (150 × 20) column; wavelength 220 nm; mobile phase: A MeCN (0.1% TFA), B water (0.1% TFA); flow rate: 25 mL/min; injection volume: 2 mL; run time: 30 min; equilibration: 5 min.

4.5.1. General procedure for preparation of compound 18

To a solution of compound **17** (20.0 g) and ammonium acetate (25.5 g) in AcOH (265 mL) was added CH₃NO₂ (21 mL). The mixture

Table 6IC₅₀ values of synthesized compounds with high inhibitory activity


Compound	X	R ₁	R ₂	R ₃	IC ₅₀ (nM)
25a	NH ₂	CH ₃	H	<i>m</i> -NH ₂ -C ₆ H ₄	50.8
23a	H	CH ₃	H	<i>m</i> -NH ₂ -C ₆ H ₄	61.8
25b	NH ₂	CH ₃	H	<i>p</i> -NH ₂ -C ₆ H ₄	68.3
26a	H	CH ₃	H	<i>m</i> -NH(CH ₂) ₂ NH ₂ -C ₆ H ₄	89.1
26b	H	CH ₃	H	<i>p</i> -NH(CH ₂) ₂ NH ₂ -C ₆ H ₄	149.3
23b	H	CH ₃	H	<i>p</i> -NH ₂ -C ₆ H ₄	367.6
32b	H	CH ₃	<i>p</i> -NH(CH ₂) ₂ NH ₂ -C ₆ H ₄	CH ₃	415.3
32a	H	CH ₃	<i>m</i> -NH(CH ₂) ₂ NH ₂ -C ₆ H ₄	CH ₃	604.5

**Figure 5.** Predicted binding mode of compound **25a**.

was stirred at 115 °C for 2 h. After cooling down to room temperature, the solvent was removed under reduced pressure. The resulting mixture was washed with saturated NaHCO₃ and extracted with CH₂Cl₂. The combined organic layers were concentrated and purified by column chromatography to give the product (10.0 g).

4.5.2. General procedure for preparation of compound **19**

To a solution of ethyl 3-oxobutanoate (6.2 mL) in methanol (60 mL) at 0 °C was added sodium methoxide (560 mg). Compound **18** (9.0 g) was then added to the mixture at the same temperature. The mixture was stirred at room temperature for 1 h. Subsequently, 5 M NH₃/methanol solution (200 mL) was poured into the mixture and the resulting mixture was stirred at room temperature overnight. The mixture was poured into AcOEt. After NH₃ was given off, the solvent was removed under reduced pressure. The resulting mixture was washed with water and extracted with

Table 7Kinase panel assay of compound **23a**. Compounds **23a** (4 μM) were tested against 21 tyrosine kinases

Kinase	% Inhibition at 4 μM
ALK	97.6
EML4-ALK	98.7
NPM1-ALK	93.8
ABL	5.1
CSK	−6.5
EGFR	−2.3
EPHA2	4.9
EPHB4	5.1
FGFR1	62.7
FLT3	98.1
IGF-1R	91.2
INSR	97.2
ITK	34.2
JAK3	9.0
KDR	81.7
LCK	69.8
MET	12.7
PDGFRα	56.6
PYK2	86.5
SRC	39.0
SYK	58.5
TIE2	14.0
TRKA	99.4
TYRO3	100.5

AcOEt. The combined organic layers were concentrated and purified by column chromatography to give the product (4.2 g).

4.5.3. General procedure for preparation of compound **20**

To compound **19** (1.0 g) was added TFA (10 mL) at 0 °C. The mixture was then stirred and refluxed overnight. After cooling down to room temperature, the solvent was removed under reduced pressure. The residue was washed with saturated NaHCO₃ and extracted with AcOEt. The combined organic layers were concentrated to give the crude product. The crude product was recrystallized by ethyl ether and petroleum ether to afford the product (600 mg) as a red solid.

4.5.4. General procedure for preparation of compound **21**

To DMF (2.7 mL) under N₂ at 0 °C was added POCl₃ (3.2 mL) drop wise over 5 min. The cooling bath was removed, and after 15 min 1,2-dichloroethane (15 mL) was added. The reaction mixture was again cooled to 0 °C and a solution of compound **20** (3.0 g) in 1,2-dichloroethane (15 mL) was added drop wise over 15 min. The reaction was heated to reflux for 15 min, and then cooled to room temperature. A solution of sodium acetate (14.4 g) in water (45 mL) was added slowly to the reaction mixture, and the resulting mixture was again heated to reflux for 20 min. After the reaction mixture was cooled to room temperature, it was diluted with CH₂Cl₂, and the aqueous phase was washed with CH₂Cl₂ (2 × 50 mL). The combined organic fractions were washed with saturated NaHCO₃, dried (Na₂SO₄), and concentrated in vacuo. The crude material was purified by chromatography on silica gel to provide 1.1 g of desired compound **21**.

4.5.5. General procedure for preparation of compound **22**

To a solution of compound **21** (5.5 g) and indolin-2-one (2.5 g) in methanol (50 mL) was added piperidine (0.5 mL). The mixture was stirred at 65 °C overnight. After cooling down to room temperature, the resulting solid gave desired compound **22** (5.6 g) as a yellow solid.

4.5.6. 3-[[3-(3-Aminophenyl)-5-methyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (23a)

To a solution of compound **22** (550 mg) in THF (60 mL) was added 10% Pd/C catalyst (90 mg). The black mixture was then stirred overnight at room temperature under an atmosphere of hydrogen. After the removal of catalyst, the filtrate was concentrated and purified by flash silica chromatography to give compound **23a** (260 mg) as a red solid. ¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.59 (s, 1H), 10.87 (s, 1H), 7.48 (s, 1H), 7.30–7.32 (d, 2H), 7.08–7.15 (m, 2H), 6.87–6.98 (m, 2H), 6.70–6.71 (d, 1H), 6.57–6.62 (m, 1H), 6.23–6.24 (d, 1H), 2.38 (s, 3H); HRMS: calcd for C₂₀H₁₈N₃O (M+H)⁺, 316.14499; found: 316.14822.

4.5.7. 3-[[3-(4-Aminophenyl)-5-methyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (23b)

¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.51 (s, 1H), 10.79 (s, 1H), 7.39 (s, 1H), 7.28–7.31 (d, 1H), 7.04–7.15 (m, 3H), 6.64–6.67 (d, 2H), 6.87–6.94 (m, 2H), 6.17–6.17 (d, 1H), 5.24 (s, 2H), 2.36 (s, 3H); HRMS: calcd for C₂₀H₁₈N₃O (M+H)⁺, 316.14499; found: 316.14889.

4.5.8. General procedure for preparation of compound 24

To a solution of compound **21** (320 mg) and 5-nitroindolin-2-one (200 mg) in methanol (5 mL) was added piperidine (0.05 mL). The mixture was stirred at 65 °C overnight. After cooling down to room temperature, the resulting solid gave desired compound **24** (400 mg) as a yellow solid.

4.5.9. 5-Amino-3-[[3-(3-aminophenyl)-5-methyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (25a)

¹H NMR (DMSO-*d*₆, 300 MHz): δ 2.39 (s, 3H), 4.70 (s, 2H), 5.22 (s, 2H), 6.19 (d, 1H), 6.37–6.40 (m, 1H), 6.57–6.60 (m, 4H), 6.67 (s, 1H), 7.13 (t, 1H), 7.31 (s, 1H), 10.45 (s, 1H), 13.72 (s, 1H); HRMS: calcd for C₂₀H₁₉N₄O (M+H)⁺, 331.15589; found: 331.15984.

4.5.10. 5-Amino-3-[[3-(4-aminophenyl)-5-methyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (25b)

¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.57–13.58 (d, 1H), 10.37 (s, 1H), 7.20 (s, 1H), 7.08–7.11 (m, 2H), 6.64–6.66 (d, 2H), 6.53–6.56 (m, 2H), 6.31–6.35 (m, 1H), 6.11–6.12 (d, 1H), 5.26 (s, 2H), 4.66 (s, 2H), 2.34 (s, 3H); HRMS: calcd for C₂₀H₁₉N₄O (M+H)⁺, 331.15589; found: 331.15814.

4.5.11. General procedure for preparation of compound 26

Compound **23** (300 mg), 2-bromoethylamine HBr salt (298 mg) and K₂CO₃ (66 mg) was mixed with DMF (15 mL). The mixture was stirred at 100 °C for 3.5 h. After the removal of the solid, the filtrate was concentrated. The residue was washed with saturated water and extracted with CH₂Cl₂. The combined organic layers were concentrated and purified by column chromatography to give the product (60 mg).

4.5.12. 3-[[3-[3-[(2-Aminoethyl)amino]phenyl]]-5-methyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (26a)

¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.63 (s, 1H), 10.87 (s, 1H), 7.68 (s, 2H), 7.47 (s, 1H), 7.29–7.32 (d, 1H), 7.19–7.25 (m, 1H), 7.07–7.13 (m, 1H), 6.92–6.97 (m, 1H), 6.86–6.89 (d, 1H), 6.68–6.70 (m, 2H), 6.61–6.64 (m, 1H), 6.25–6.26 (d, 1H), 5.88–5.90 (m, 1H), 3.28–3.33 (m, 2H), 2.95–2.99 (m, 2H), 2.39 (s, 3H); HRMS: calcd for C₂₂H₂₃N₄O (M+H)⁺, 359.18719; found: 359.186.

4.5.13. 3-[[3-[4-[(2-Aminoethyl)amino]phenyl]]-5-methyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (26b)

¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.53 (s, 1H), 7.40 (s, 1H), 7.22–7.29 (m, 3H), 7.04–7.07 (m, 1H), 6.86–6.95 (m, 2H), 6.71–6.74 (d, 2H), 6.18–6.19 (d, 1H), 6.05 (s, 1H), 3.30–3.32 (m, 2H),

2.93–2.97 (m, 2H), 2.36 (s, 3H); HRMS: calcd for C₂₂H₂₃N₄O (M+H)⁺, 359.18719; found: 359.18838.

4.5.14. Preparation of compound 28

To a solution of 2-formyl-3,5-dimethylpyrrole (compound **27**, 7.17 g) and indolin-2-one (5.0 g) in methanol (100 mL) was added piperidine (1.0 mL). The mixture was stirred at 65 °C overnight. After cooling down to room temperature, the resulting solid gave desired compound **29** (10.4 g) as a yellow solid.

4.5.15. Preparation of compound 29

Compound **28** (1.0 g), NBS (0.785 g) and benzoyl peroxide (25 mg) were mixed with CCl₄ (100 mL). The mixture was stirred at room temperature overnight. The resulting solid afforded a crude product. The crude product was recrystallized by methanol to give the product (1.1 g) as a yellow solid.

4.5.16. General procedure for preparation of compound 30

A mixture of compound **29** (1.0 g), 4-nitrophenylboronic acid (0.79 g), K₂CO₃ (1.31 g), Pd(dppf)Cl₂ (258 mg), H₂O (4.5 mL), and dioxane (10 mL) was heated to 105 °C for 1 h. After cooling down to room temperature, water was added. The mixture was then extracted with CH₂Cl₂ and MeOH. The combined organic layer was dried over anhydrous sodium sulfate and concentrated. The residue was purified by column chromatography to give the product (850 mg) as a yellow solid.

4.5.17. 3-[[4-(3-Aminophenyl)-3,5-dimethyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (31a)

To a solution of compound **30** (700 mg) in THF (80 mL) was added 10% Pd/C catalyst (120 mg). The black mixture was then stirred overnight at room temperature under an atmosphere of hydrogen. After the removal of catalyst, the filtrate was concentrated and purified by flash silica chromatography to give compound **1** (500 mg) as a red solid. ¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.59 (s, 1H), 10.80 (s, 1H), 7.71–7.73 (d, 1H), 7.62 (s, 1H), 7.04–7.08 (m, 2H), 6.95–6.96 (m, 1H), 6.85–6.87 (m, 1H), 6.40–6.49 (m, 3H), 5.04 (s, 2H), 2.26–2.30 (d, 6H); HRMS: calcd for C₂₁H₂₀N₃O (M+H)⁺, 330.16064; found: 330.16533.

4.5.18. 3-[[4-(4-Aminophenyl)-3,5-dimethyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (31b)

¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.56 (s, 1H), 10.74 (s, 1H), 7.68–7.71 (d, 1H), 7.59 (s, 1H), 7.04–7.07 (m, 1H), 6.92–6.97 (m, 3H), 6.84–6.86 (d, 1H), 6.59–6.62 (d, 2H), 5.04 (s, 2H), 2.24–2.28 (d, 6H); HRMS: calcd for C₂₁H₂₀N₃O (M+H)⁺, 330.16064; found: 330.15895.

4.5.19. 3-[[4-[3-[(2-Aminoethyl)amino]phenyl]]-3,5-dimethyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (32a)

Compound **31a** (470 mg), 2-bromoethylamine HBr salt (3.2 g) and Et₃N (2.16 mL) was mixed with isopropanol (50 mL). The mixture was stirred and refluxed for 24 h. The solvent was removed under reduced pressure. The residue was purified by column chromatography to give the product (410 mg). ¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.64 (s, 1H), 10.82 (s, 1H), 7.73–7.75 (m, 3H), 7.65 (s, 1H), 7.08–7.20 (m, 2H), 6.98–7.00 (m, 1H), 6.87–6.89 (d, 1H), 6.53–6.57 (m, 3H), 5.75–5.80 (m, 1H), 3.29–3.38 (m, 2H), 2.97–2.99 (m, 2H), 2.29–2.33 (d, 6H); HRMS: calcd for C₂₃H₂₅N₄O (M+H)⁺, 373.20284; found: 373.20285.

4.5.20. 3-[[4-[4-[(2-Aminoethyl)amino]phenyl]]-3,5-dimethyl-1H-pyrrol-2-yl]methylene]-1,3-dihydro-2H-indol-2-one (32b)

¹H NMR (DMSO-*d*₆, 300 MHz): δ 13.61 (s, 1H), 10.79 (s, 1H), 7.71–7.80 (m, 4H), 7.62 (s, 1H), 7.06–7.12 (m, 3H), 6.94–6.99 (m, 1H), 6.86–6.89 (d, 1H), 6.67–6.70 (d, 2H), 5.78 (s, 1H), 3.32–3.41

(m, 2H), 2.97–3.01 (m, 2H), 2.26–2.30 (d, 6H); HRMS: calcd for $C_{23}H_{25}N_4O$ (M+H)⁺, 373.20284; found: 373.20581.

Acknowledgements

This work was supported by the Target Protein Research Programs of the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References and notes

- Morris, S. W.; Kirstein, M. N.; Valentine, M. B.; Dittmer, K. G.; Shapiro, D. N.; Saltman, D. L.; Look, A. T. *Science* **1994**, *263*, 1281.
- Galkin, A. V.; Melnick, J. S.; Kim, S.; Hood, T. L.; Li, N.; Li, L.; Xia, G.; Steensma, R.; Chopiuk, G.; Jiang, J.; Wan, Y.; Ding, P.; Liu, Y.; Sun, F.; Schultz, P. G.; Gray, N. S.; Warmuth, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 270.
- Soda, M.; Choi, Y. L.; Enomoto, M.; Takada, S.; Yamashita, Y.; Ishikawa, S.; Fujiwara, S.; Watanabe, H.; Kurashina, K.; Hatanaka, H.; Bando, M.; Ohno, S.; Ishikawa, Y.; Aburatani, H.; Niki, T.; Sohara, Y.; Sugiyama, Y.; Mano, H. *Nature* **2007**, *448*, 561.
- Zou, H. Y.; Li, Q.; Lee, J. H.; Arango, M. E.; McDonnell, S. R.; Yamazaki, Shinji; Koudriakova, T. B.; Alton, G.; Cui, J. J.; Kung, P.-P.; Nambu, M. D.; Los, G.; Bender, S. L.; Mroczkowski, B.; Christensen, J. G. *Cancer Res.* **2007**, *67*, 4408.
- Okamoto, M.; Takayama, K.; Shimizu, T.; Ishida, K.; Takahashi, O.; Furuya, T. *J. Med. Chem.* **2009**, *52*, 7323.
- Okamoto, M.; Saito, N.; Kojima, H.; Okabe, T.; Takeda, K.; Ichijo, H.; Furuya, T.; Nagano, T. *Bioorg. Med. Chem.* **2011**, *19*, 486.
- Chemical Biology Research Initiative (CBRI). http://www.cbri.u-tokyo.ac.jp/index_e.html.
- Hubbard, S. R. *EMBO J.* **1997**, *16*, 5572.
- Okamoto, M.; Masuda, Y.; Muroya, A.; Yasuno, K.; Takahashi, O.; Furuya, T. *Chem. Pharm. Bull.* **2010**, *58*, 1655.
- Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
- (a) Lee, C. C.; Jia, Y.; Li, N.; Sun, X.; Ng, K.; Ambing, E.; Gao, M.-Y.; Hua, S.; Chen, C.; Kim, S.; Michellys, P.-Y.; Lesley, S. A.; Harris, J. L.; Spraggon, G. *Biochem. J.* **2010**, *430*, 425 (PDB code: 3LCT, 3L9P, 3LCS); (b) Bossi, R. T.; Saccardo, M. B.; Ardini, E.; Menichincheri, M.; Rusconi, L.; Magnaghi, P.; Orsini, P.; Avanzi, N.; Borgia, A. L.; Nesi, M.; Bandiera, T.; Fogliatto, G.; Bertrand, J. A. *Biochemistry* **2010**, *49*, 6813 (PDB code: 2XBA, 2XB7); (c) Mctigue, M.; Deng, Y.; Liu, W.; Brooun, A.; Timofeevski, S.; Marrone, T.; Cui, J. (PDB code: 2XP2).
- Basic Local Alignment Search Tool (BLAST). <http://blast.ncbi.nlm.nih.gov/Blast.cgi>.
- Gunby, R. H.; Ahmed, S.; Sottocornola, R.; Gasser, M.; Redaelli, S.; Mologni, L.; Tartari, C. J.; Belloni, V.; Gambacorti-Passerini, C.; Scapozza, L. *J. Med. Chem.* **2006**, *49*, 5759.
- Li, R.; Xue, L.; Zhu, T.; Jiang, Q.; Cui, X.; Yan, Z.; McGee, D.; Wang, J.; Gantla, V. R.; Pickens, J. C.; McGrath, D.; Chucholowski, A.; Morris, S. W.; Webb, T. R. *J. Med. Chem.* **2006**, *49*, 1006.
- Molecular Operating Environment (MOE 2008.1001), Chemical Computing Group Inc., 1010 Sherbrooke St. W, Suite 910, Montreal, Quebec, Canada H3A 2R7.
- Exelixis, Inc., WO 2005009389 A2.