Computational and Structural Biotechnology Journal 18 (2020) 3936–3946

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

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

Structure-based discovery of neoandrographolide as a novel inhibitor of Rab5 to suppress cancer growth

Jing Zhang [a,b,1], Yue Sun [a,1], Li-Ye Zhong [a], Nan-Nan Yu [a], Lan Ouyang [a], Run-Dong Fang [a], Yang Wang [a,*], Qing-Yu He [a,b,*]

[a] MOE Key Laboratory of Tumor Molecular Biology and Key Laboratory of Functional Protein Research of Guangdong Higher Education Institutes, Institute of Life and Health Engineering, College of Life Science and Technology, Jinan University, Guangzhou 510632, China
[b] The First Affiliated Hospital, Jinan University, Guangzhou 510632, China

ARTICLE INFO

Article history:
Received 28 August 2020
Received in revised form 20 November 2020
Accepted 21 November 2020
Available online 30 November 2020

Keywords:
Virtual screening
Rab5
Neoandrographolide
Cancer therapy

ABSTRACT

Rab5 is a small GTPase that plays a crucial role in oncogenic signal transduction, which was considered as an attractive target for cancer therapy. Rapid GDP/GTP exchange in the packet of Rab5 sustains its high activity for promoting cancer progression. However, Rab5 currently remains undruggable due to the lack of specific inhibitor. Herein, we reported the discovery of a novel Rab5 inhibitor, neoandrographolide (NAP), by using high-throughput virtual screening with a natural product library containing 7459 compounds, which can occupy the surface groove of Rab5, competing with GDP/GTP for the binding. Ser34 is the most important residue in the groove of Rab5, as it forms most hydrogen-bond interactions with GDP/GTP or NAP, and *in silico* mutation of Ser34 decreased the stabilization of Rab5. Moreover, fluorescence titration experiment and isothermal titration calorimetry (ITC) assay revealed a direct binding between NAP and Rab5. Biochemical and cell-based assays showed that NAP treatment not only diminished the activity of Rab5, but also suppressed cell growth of cancer cell. This finding firstly identifies NAP as a novel inhibitor of Rab5, which directly binds with Rab5 by occupying the GDP/GTP binding groove to suppress its functions, highlighting a great potential of NAP to be developed as a chemotherapeutic agent in cancer therapy.

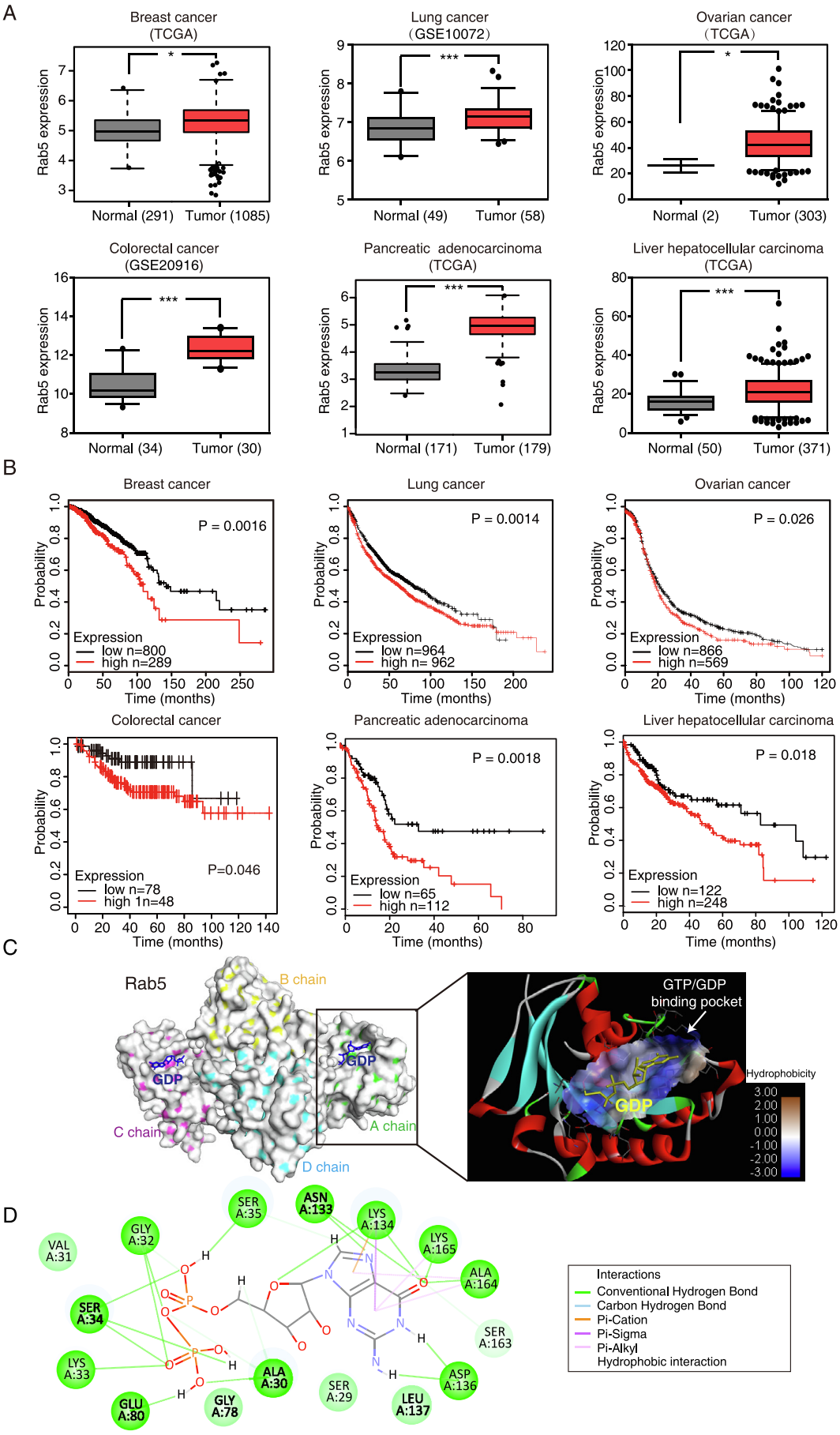
1. Introduction

Rab5, a prototypical Ras analog in brain (Rab) protein, controls intracellular membrane transport, including vesicle formation, fusion and maturation [1], as well as receptor-mediated endocytosis. As a small GTPase, Rab5 exerts its biological function by exchanging GDP- and GTP-bound states, which was promoted by specific guanine exchange factors (GEFs), such as Rin1 [2]. Activated Rab5 translocates from cytoplasma to cellular membrane to control receptor endocytosis and signaling transduction by binding with its effectors. Upon ligand binding, the receptor tyrosine kinase EGFR (epidermal growth factor receptor) activates Rab5 and undergoes a rapid internalization [3], while inactivated Rab5 represses the EGF-stimulated EGFR internalization and degradation, decreasing EGFR-MAPK/ERK signaling-mediated cell proliferation [2]. EGFR signal can be potentiated by the expression of the activated form of Rab5. Moreover, Rab5 was reported to regulate myoblast differentiation by activating IRS1-IGF-AKT-mTOR signaling during muscle regeneration [4]. In the aspect of cell mobility, Rab5 induces focal adhesion kinase-dependent activation to control cell migration [5]. It can change cell morphology from spindle to round by inhibiting E-cadherin expression, increasing proliferation, invasion, and cell migration in pancreatic cancer [6]. As such, the role of Rab5 in cancer has been extensively studied. High expression of Rab5 was observed in multiple types of human malignancies including breast and pancreatic cancers, and was correlated with tumor progression and poor prognosis [6,7]. However, the specific inhibitor of Rab5 remains unclear, so that the exploration of its specific inhibitor will provide a new insight into clinical cancer therapy.

Structure-based virtual screening is an effective strategy to identify the binding small molecules (ligands) for some undruggable proteins, and to predict their binding sites in detail based on binding energy [8,9]. Moreover, recent computational approach

* Corresponding authors at: Institute of Life and Health Engineering, College of Life Science and Technology, Jinan University, Guangzhou 510632, China.
E-mail addresses: wangyang0507@jnu.edu.cn (Y. Wang), tqyhe@email.jnu.edu.cn (Q.-Y. He).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.csbj.2020.11.033
2001-0370/


A
Breast cancer (TCGA)
Lung cancer (GSE10072)
Ovarian cancer (TCGA)
Colorectal cancer (GSE20916)
Pancreatic adenocarcinoma (TCGA)
Liver hepatocellular carcinoma (TCGA)
Rab5 expression
Normal (291) Tumor (1085)
Normal (49) Tumor (58)
Normal (2) Tumor (303)
Normal (34) Tumor (30)
Normal (171) Tumor (179)
Normal (50) Tumor (371)
B
Breast cancer
Lung cancer
Ovarian cancer
Colorectal cancer
Pancreatic adenocarcinoma
Liver hepatocellular carcinoma
Probability
Time (months)
P = 0.0016
P = 0.0014
P = 0.026
P=0.046
P = 0.0018
P = 0.018
Expression
low n=800
high n=289
low n=964
high n= 962
low n=866
high n=569
low n=78
high 1n=48
low n=65
high n=112
low n=122
high n=248
C
Rab5
B chain
GDP
GTP
C chain
A chain
D chain
GTP/GDP binding pocket
Hydrophobicity
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
D
SER A:35
ASN A:133
LYS A:134
LYS A:165
ALA A:164
SER A:163
ASP A:136
LEU A:137
SER A:29
ALA A:30
GLY A:78
GLU A:80
LYS A:33
SER A:34
GLY A:32
VAL A:31
Interactions
Conventional Hydrogen Bond
Carbon Hydrogen Bond
Pi-Cation
Pi-Sigma
Pi-Alkyl
Hydrophobic interaction

considers the pharmacokinetic parameters including absorption, distribution, metabolism, excretion and toxicity (ADMET) screening to improve the process of drug selection, serving as a promising tool for identifying potential drug candidates [10]. Natural products are an important source for the discovery of therapeutic agents [11,12], in particular, compounds from drug-food homologous plants will provide higher safety for clinical application. For example, curcumol, a guaiane-type sesquiterpenoid hemiketal isolated form *Rhizoma Curcumae*, exhibits nontoxicity but synergistic lethal effects in combination with celecoxib on non-small cell lung cancer [13]. Liensinine perchlorate, a constituent of *Nelumbo nucifera Gaertn*, exerts a significant inhibitory effect on colorectal cancer cells without causing cytotoxicity on normal colorectal epithelial cells [14]. Regarding the significant oncogenetic activities of Rab5 in carcinogenesis, the rapid activation-inactivation cycle of Rab5 in the cells remains lack of selective small-molecule inhibitors to limit this process directly.

In this study, we performed high-throughput virtual screening with a natural product library containing 7459 compounds to identify the selective Rab5 inhibitor that can compete with GDP/GTP for the binding groove on the surface of Rab5. Neoandrographolide (NAP), a compound isolated from drug-food homologous plant Andrographis paniculate, with higher docking score and good ADMET prediction, was identified. Using a series of *in silico* and cell experimental validations, we demonstrated that NAP is a novel Rab5 inhibitor that occupies the GTPase groove to limit the Rab5 activation, resulting in the subsequent blockage of EGFR degradation and ERK signaling-mediated cell proliferation. NAP may be further explored to be an anti-tumor drug targeting Rab5 for the clinical cancer therapy.

2. Materials and methods

2.1. Gene expression and survival data analysis

Rab5 gene expression and survival data of cohorts of patients with different types of cancers were downloaded from the GEO (Gene Expression Omnibus) or TCGA (The Cancer Genome Atlas) database. The expression values of Rab5 and clinical data from the data matrices were extracted for analysis, and the expression values were further divided into positive and negative parts using the median expression level as the cutoff point for Kaplan-Meier survival analyses.

2.2. Chemicals and antibodies

Neoandrographolide (NAP, CAS No.: 27215-14-1) was obtained from TargetMol (Shanghai, China). Antibodies: Rab5, GST, EGFR, Rin1, β-actin and the HRP-conjugated goat anti-rabbit/mouse secondary antibody were purchased from Proteintech (Wuhan, Hubei, China). ERK1/2, Phospho-ERK1-T202/Y204 + ERK2-T185/Y187 pAb was purchased from Abclonal (Wuhan, Hubei, China).

2.3. Virtual screening

Virtual screening was performed by using Discovery Studio 4.5 software (DS4.5, Accelrys Inc., San Diego, CA, USA) on a natural product library including 7459 compounds constructed by merging TargetMol and Selleck companies. The crystal structure of Rab5 (1TU4) was downloaded from the Protein Data Bank (PDB; http://www.rcsb.org/pdb). The docking pocket in Rab5 was created according to the residues binding with GDP using DS4.5. For protein preparation of 1TU4, all water molecules and ions were removed, the box radius for docking is xyz: 51.17, 4.61 and 50.90, Radius: 9.89. ADMET tool was used to select the compounds with optimal absorption, distribution, metabolism, excretion, and toxicity properties at the early stage of virtual screening. ADMET estimated pharmacokinetic properties including: HIA (human intestinal absorption), BBB (blood–brain barrier), PPB (plasma protein binding), Cytochrome P450 2D6 binding (CYP2D6), Aqueous solubility, and Toxicity (hepatotoxicity). The ADMET-filtered natural products were subjected to high-throughput virtual screening LibDock docking, and the top-ranked 50 molecules were subjected to TOPKAT (Toxicity Prediction) assay, to filter out those compounds with toxicity, mutagenicity and carcinogenicity.

2.4. Molecular dynamics (MD)

MD simulations were performed using trajectory analysis tools (LARMD webserver) to calculate the root mean square deviation (RMSD) and the atomic RMS fluctuation (RMSF)

2.5. Ligand poses analysis and mutation energy calculation

Non-bond interactions formed between ligands and receptor (Rab5) were calculated by the algorithm of ligand poses analysis. The docking scores based on CDOCKER interaction energy were used to rank poses, and the top ten poses for each ligand (NAP, GDP, GTP) were selected to analyze the possible significant non-bond interactions.

To evaluate the effect of each mutation on protein stability, the algorithm for mutation energy calculation was performed with DS4.5. The effect of mutation and the corresponding mutation energy terms are defined as follows: Stabilizing, mutation energy is less than −0.5 kcal/mol; Neutral, between −0.5 and 0.5 kcal/mol; Destabilizing, greater than 0.5 kcal/mol.

2.6. Protein purification and isothermal titration calorimetry (ITC) assay

Human Rab5, Rab7 and Rab11 were constructed into the GST fusion vector pGEX-4 T-1(GE Healthcare, Piscataway, USA), and the proteins were expressed by *Escherichia coli* BL21. When the optical density at 600 nm (OD600) reached at 0.6, 0.5 mM of isopropyl-β-D-thiogalactoside (IPTG) was added to induce protein expression. After additional 6 h of induction, BL21 was harvested in binding buffer (PBS: 10 mM Na_2PO_4, 1.8 mM KH_2PO_4, 140 mM NaCl, 2.7 mM KCl, pH = 7.3) and the GST-fused proteins were purified by glutathione-sepharose 4B resin (Beyotime, Jiangsu, China) according to the manufacturer's instructions. The GST-fused proteins were eluted by elution buffer (10 mM GSH, 50 mM Tris-HCl, pH = 8.0), and then digested by PreScission Protease (Beyotime) in digestion buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH = 7.5) overnight at 4 °C. The cleaved GST-tag was removed by glutathione-sepharose 4B resin.

Fig. 1. The oncogenic role of Rab5 and its structure in binding with GDP. (A) Pan-cancer expression landscape of Rab5 according to GEPIA. Tumor tissue (Red); Normal, paired normal tissue (Gray). (B) Kaplan-Meier survival analysis of Rab5 in 6 types of cancer using the GEO and TCGA databases, showing that patients with high Rab5 expression were correlated with poor prognosis. (C) The 3D structure of Rab5 (1T4U) showing that a tetramer was formed by 4 subunits of Rab5, each bound a GDP molecule. In the binding pocket of GDP, blue represents relevant residues with low hydrophobicity. (D) 2D diagram showing the interactions between GDP and Rab5 residues, the formed bonds were presented by indicated colors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

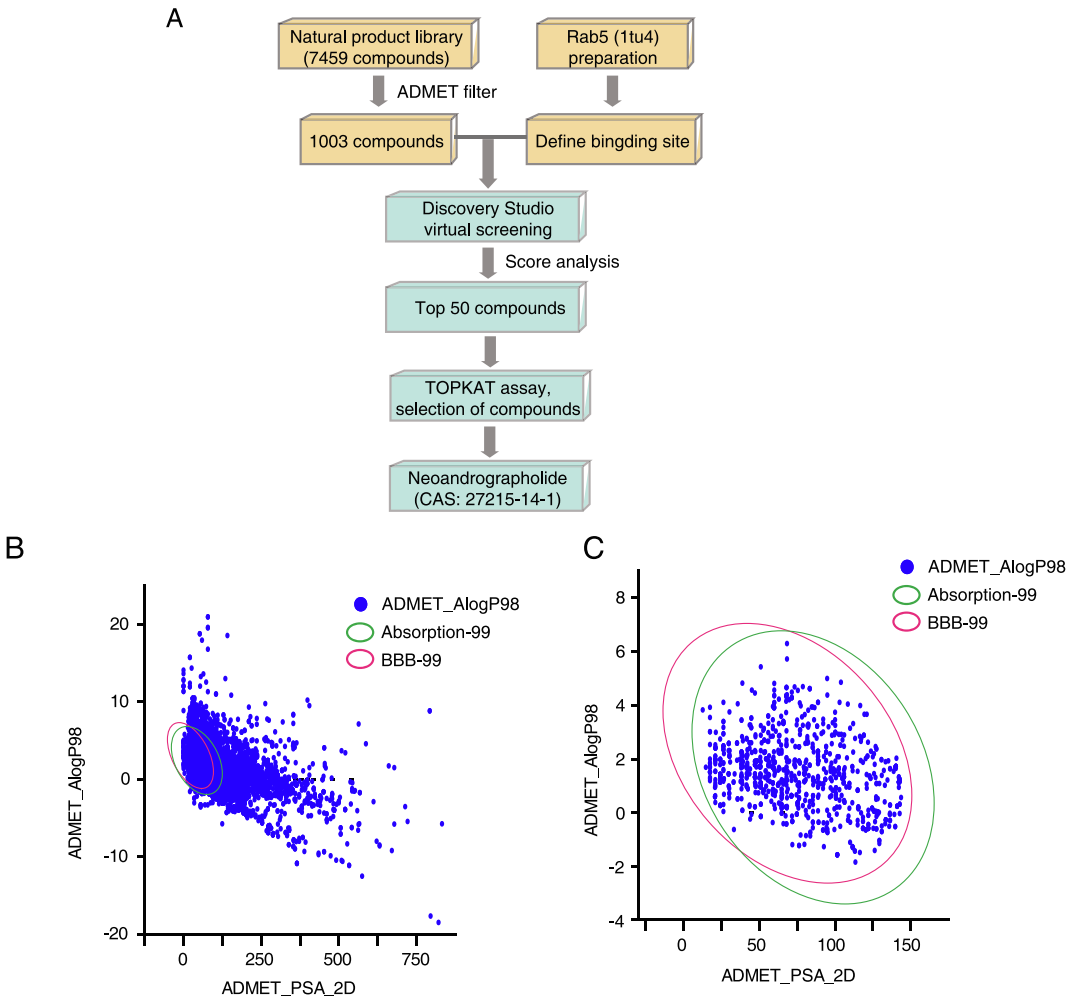


Fig. 2. Virtual screening of Rab5 inhibitor. (A) Scheme depicting the approach for the identification of Rab5 inhibitors. (B) Plot of PSA and AlogP98 for 7459 compounds in the natural product library showing 99% confidence limit ellipses corresponding to the human intestinal absorption models (absorption) and blood–brain barrier (BBB) (Green ellipse depicts 99% of absorption, red ellipse depicts 99% of BBB). (C) Plot of PSA and AlogP98 for the 1003 compounds generated by applying ADMET filter. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

For the binding of recombinant Rab5, Rab7 or Rab11 with NAP, ITC experiments were performed by using a MicroCalorimeter Auto-ITC 200 (Malvern, UK). In brief, recombinant Rab5, Rab7 or Rab11 (20 μM in PBS) was titrated with NAP (200 μM in PBS) by a syringe. Titration parameters: 2 μL of NAP solution were injected into the 200 μL protein sample cell during each titration with 18 injections. The delay time between injections was set at 3 min. Stirring speed was set at 750 rpm. The background heat effect was subtracted by addition of NAP alone to the buffer.

2.7. Measurement of Rab5 activity

The activity of Rab5 was determined by the Rab5-binding domain of GST-fusion protein (R5BD-GST) pull-down assay. R5BD-GST was expressed by pGEX-4 T-1 in *Escherichia coli* BL21 induced by 6 h of IPTG, and purified by glutathione-sepharose 4B resin. Cell lysates (1500 μg) with different treatment were incubated with 30 μL of R5BD-GST bound beads for 10 min at 4 ℃ under rotation. The complexes were subsequently rinsed with binding buffer, and then subjected to SDS-PAGE for immunoblotted with anti-Rab5 antibody.

2.8. Fluorescence titration

Recombination Rab5, Rab7 or Rab11 protein was diluted to 1 μM in PBS and the NAP stock solution was dissolved to 0.6 mM in PBS. Aliquots of NAP were gradually added to 1.5 mL Rab protein solution and incubated for 1 min at 37 °C before each fluorescence spectra measurement (Hitachi F7000). The excitation wavelength was 280 nm and the emission wavelength started from 290 nm to 500 nm. The fluorescence measured at 330 nm was extracted to produce the titration curve. The titration curve was fitted with the Hill plot equation $y = V_{max} \times x_n/(k_n + x_n)$ in Origin 8.5 software [15], and the dissociation constant (K_d) was calculated [16].

2.9. Western blotting

Western blot assay was performed as referenced [17]. The primary and secondary antibodies used in this study have been described above. The signals were visualized with the ECL reagent (BIO-RAD), and the images were obtained using a Tannon 5200-Multi (Tanon Science & Technology Co. Ltd, Shanghai, China).

2.10. Cell culture and plasmids transfection

Lung cancer cell line A549 and epithelium cell 293 T (American Type Culture Collection, Manassas, VA, USA) were cultured in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% (v/v) fetal bovine serum and 1% (v/v) penicillin–streptomycin, at 37 ℃ in 5% CO_2. Stable cell lines with expression of wild-type Rab5 or Rab5 mutant S34N were established based on retrovirus vectors. For infection, 293 T cells were transfected with lentiviral plasmids and virus skeleton vectors using Lipofectamine 3000 (Thermo Fisher Scientific). A549 cells were infected with virus-containing culture media, and selected with puromycin. The cells with stable overexpression of wild-type or mutant Rab5 were established and maintained at least two weeks before experiments.

2.11. Cellular assays

WST-1 assay (Beyotime) was used to determine the cytotoxicity of NAP. In brief, A549 cells (2×10^3) were seeded in 96-well plate and treated with increasing concentration of NAP (up to 400 μM) for 48 h. WST-1 was added and the plate was read after 1.5 h on an automated microplate spectrophotometer (BioTek Instruments, Vermont, USA) at 450 nm according to the manufacturer's instructions.

Colony-formation assay was performed to detect long-term survival of A549 cells in NAP treatment. 400 cells seeded in 12-well plates were treated with indicated concentration of NAP for 15 days. The plates were washed twice with PBS and fixed with methanol for 10 min at room temperature and then stained with 1% crystal violet for 5 min. The number of colonies was counted using ImageJ software (version 1.44I).

Cellular uptake assay was performed by using FITC-Dextran (TargetMol). A549 cells were seeded onto 12-well plate and incubated with FITC-dextran (0.1 mg/mL) for 24 h, then the cells were treated with 50, 100, 150 and 200 μM of NAP for 24 h and 48 h, respectively. The fluorescence (excited at 488 nm and emitted at 530 nm) was measured with a multifunctional microplate reader Spark 10 M (Tecan, Synergy-HT).

2.12. Statistical analysis

In vitro experiments were repeated three times and the data were presented as the mean ± SD. One-way ANOVA analysis was performed using GraphPad Prism software v.5.01. Survival analysis was performed using the Kaplan-Meier method with the log-rank test by IBM SPSS Statistics v.19. The differences with $^{*}P < 0.05$, $^{**}P < 0.01$ or $^{***}P < 0.001$ were considered statistically significant.

3. Results and discussion

3.1. High expression of Rab5 correlates with poor survival in cancer patients

To demonstrate the role of Rab5 in cancer, the expressions of Rab5 in cancer and normal tissues of 6 cancer types, including breast cancer, lung cancer, ovarian cancer, colorectal cancer, pancreatic adenocarcinoma and hepatocellular carcinoma were compared by using the TCGA and GEO datasets. We found that Rab5 was highly expressed in cancer tissues, as compared to adjacent normal tissues (Fig. 1A). Moreover, patients with high Rab5 expression exhibited shorter survival time ($P < 0.05$) in these cancers (Fig. 1B). These results suggest an oncogenetic role of Rab5 and its clinical significance in pan-cancer, highlighting its potential to be developed as a therapeutic target for clinical cancer treatment.

As an enzyme, Rab5 has a binding pocket on its structural surface to trap its substrate GDP or GTP, rendering an inactive or active protein form. Since Rab5 cycles between inactive GDP-bound state and active GTP-bound state to exert its biological function [18], we proposed that small molecules directly bind to the key residues in the GDP/GTP pocket will effectively disrupt the GDP/GTP exchange and disable the function of Rab5, and thus would be potential anti-cancer therapeutic agents. We therefore

Table 1
ADMET properties of nine screened compounds.

NO.	Compounds	Canonical SMILES	Molecular formula	Molecular weight
1	Neoandrographolide	CC1(CCCC2(C1CCC(=C)C2CCC3 = CCOC3 = O)C)COC4C(C(C(C(O4)CO)O)O)O	C26H40O8	480.598
2	SCHEMBL4283398	CC1CC2OC(=O)C(CN3CCN(CC3)C(=O)c3ccccc3)C2C(O)C2(C)C(CC(OC(C) = O)C12)OC(C) = O	C30H40N2O8	556.656
3	Cymarol	C12CCC3(O)CC(CCC3(CO)C1CCC1(C)C(CCC12O)C1 = CC(=O)OC1)OC1CC(OC)C(O)C(C)O1	C30H46O9	550.689
4	Asperglaucide	CC(=O)OCC(Cc1ccccc1)NC(=O)C(Cc1ccccc1)NC(=O)c1ccccc1	C27H28N2O4	444.531
5	AKOS037491052	C12CCC3CC(CCC3(C)C1CCC1(C)C(C(CC12O)OC(=O)C[N +](C)(C)C)C1 = CC(=O)OC1)OC(=O)C[N +](C)(C)C	C33H54N2O7	590.801
6	/	C1(CC(COC(C) = O)OC(C) = O)CCC2C3CCC4CC(CCC4(C)C3CCC12C)OC(C) = O	C28H44O6	476.654
7	Misoprostol	C1(CC(C(C1CCCCCCC(OC) = O)C = CCC(CCCC)(C)O)O) = O	C22H38O5	382.534
8	Ticarcillin sodium	CC1(C(N2C(S1)C(C2 = O)NC(=O)C(C3 = CSC = C3)C(=O)[O-])C(=O)[O-])C.[Na +].[Na +]	C15H16N2Na2O6S2	430.407

NO.	LibDock Score	T-M[a]	T-C[b]	A[c]	D[d]	M[e]	E[f]	T[g]	AlogP98[H]	PSA_2D[i]	Rat Oral LD50 (g/kg)	Chronic LOAEL (mg/kg)
1	133.943	NA	NA	1	4	NA	NA	NA	2.378	127.353	3.8	142.3
2	130.514	NA	NA	0	4	NA	NA	NA	1.281	123.513	2.1	0.09
3	128.332	NA	NA	1	4	NA	NA	NA	1.635	136.283	10	186.2
4	125.647	NA	NA	0	2	NA	NA	NA	3.911	86.452	0.6	127.4
5	124.588	NA	NA	0	3	NA	NA	NA	0.146	99.508	10	0.21
6	122.013	NA	NA	0	1	NA	NA	NA	4.698	78.692	1.8	12.1
7	121.662	NA	NA	0	2	NA	NA	NA	3.94	85.162	6.9	173.4
8	117.631	NA	NA	1	4	NA	NA	NA	0.027	126.996	10	2.6

[H]AlogP98: Predicted octanol/water.
[a] Toxicity: Mutagenicity.
[b] Toxicity: Carcinogenicity.
[c] Absorption: Intestinal absorption.
[d] Distribution: Aqueous solubility and Blood-brain barrier penetration.
[e] Metabolism: CYPA2D6.
[f] Excretion: Plasma protein binding.
[g] Toxicity: Hepatotoxicity.
[i] PSA: polar surface area, 2D: two-dimensional.

performed high-throughput virtual screening to search for the small molecules that can occupy the active pocket of Rab5.

Firstly, we analyzed the 3D structure of Rab5 (Protein Data Bank ID: 1TU4) and observed a tetramer of Rab5 formed by four adjacent subunits with the GDP binding pocket on the surface individually (Fig. 1C). Residues Ala30, Gly32, Lys33, Ser34, Ser35, Glu80, Asn133, Lys134, Asp136, Ala164 and Lys165 on Rab5 were found to form hydrogen bonds with GDP. Furthermore, hydrophobic interactions, Pi-Sigma, Pi-Cation and Pi-Alky were found between Rab5 and GDP (Fig. 1D). To screen for Rab5 inhibitors, the putative inhibitor binding sites on Rab5 were defined according to the GDP binding pocket that is constructed by the 17 residues shown in Fig. 1D.

3.2. Virtual screening for Rab5 inhibitors

In the basis of the aforementioned binding pocket on Rab5, we performed *in silico* screening to identify Rab5 inhibitors with high safety property and high docking score (Fig. 2A). We thus established a large chemical library including 7459 natural products, and filtered the chemicals by using ADMET tool to evaluate their absorption, distribution, metabolism, excretion, and toxicity properties. Based on PSA (polar surface area) and ALogP98 criteria, 4949 compounds were found to meet the set criteria at both 99% confidence limit ellipses for the HIA and BBB penetration models (Fig. 2B).

Among the 4949 compounds, 1003 compounds were found to have good solubility, good bioavailability, non-inhibition on CYP2D6 and non-toxicity (Fig. 2C), and thus were retained for further docking filtration. Libdock was sequentially applied to evaluate the binding capacity of the 1003 ADMET-filtered compounds, and the top-ranked 50 compounds were selected for TOPKAT assay to guarantee high safety. After the filtration assay, 8 compounds were found to have non-mutagenicity and non-carcinogenicity (Table 1). Among them, NAP with the highest docking score was identified as the most promising inhibitor against Rab5, and thus was chosen for further validation.

3.3. Binding mode of NAP in Rab5

The putative Rab5 inhibitor NAP is a natural diterpenoid compound extracted from the leaves of *andrographis paniculate,* a drug-food homologous plant that is widely used for disease prevention in daily diet (Fig. 3A). The binding mode of NAP with

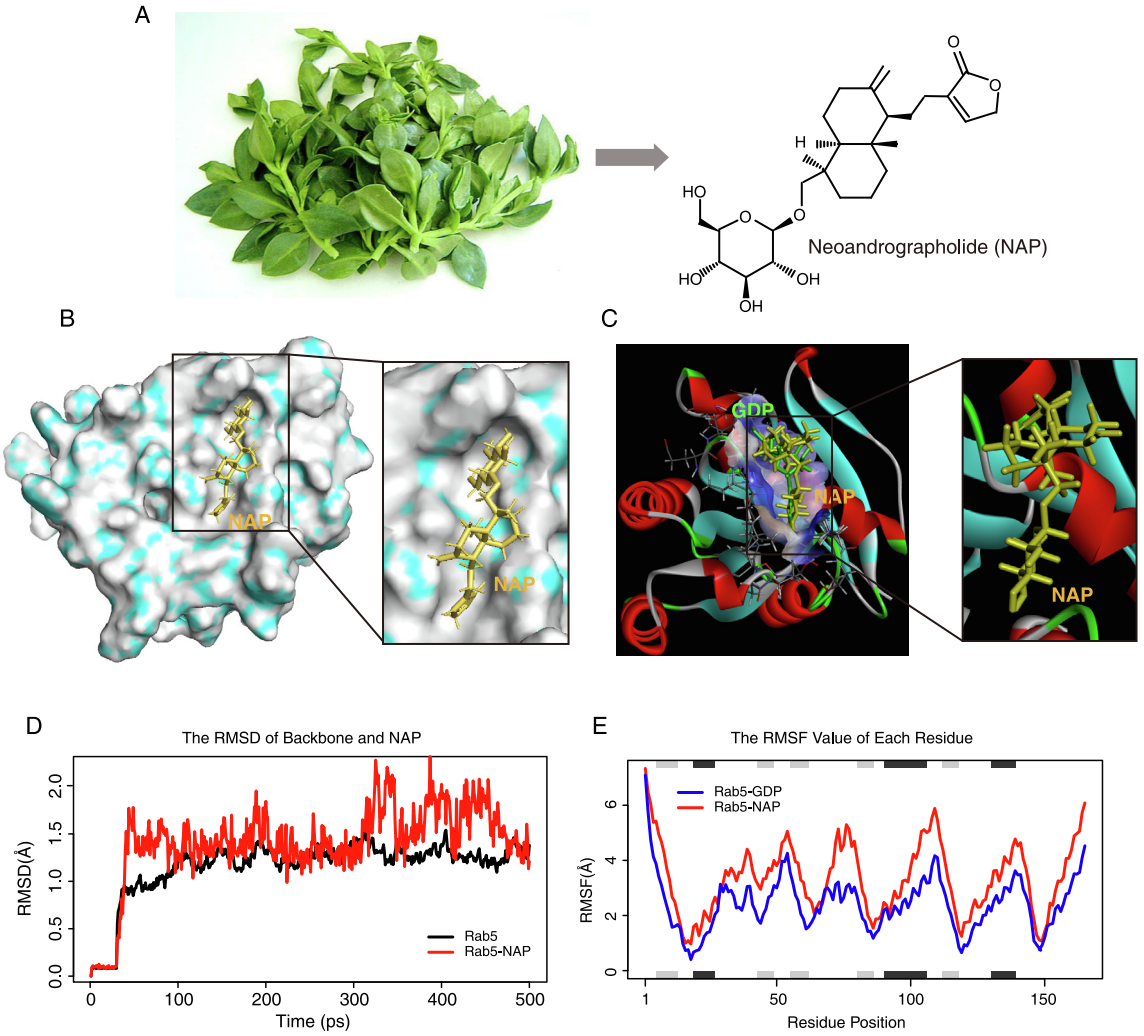


Fig. 3. NAP occupies GDP binding pocket on Rab5. (A) The chemical structure of NAP, a diterpenoid isolated from herbal *andrographitis*. (B) Binding pose depiction of potential interaction mode of Rab5 and NAP. (C) Superimposition of the docked pose of NAP (yellow) and GDP (green) in Rab5. (D) RMSD plots showing the average distance between the atoms of Rab5-NAP (red) and Rab5 backbone (black). (E) The RMSF plots showing the fluctuation of the Rab5 residues binding with NAP (red) as compared to the GDP (blue) across the production trajectory. Helices black, strands gray and loops, white. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Rab5 was similar to GTP or GDP on Rab5, NAP replaced the substrate to occupy the binding groove on the surface (Fig. 3B and 3C). We performed MD simulation to measure the stability of the Rab5-NAP complex, and found that the binding free energies (ΔPB/ΔGB) [19] for Rab5-NAP were ΔPB −27.89 kcal/mol and ΔGB −10.17 kcal/mol. The report plots the RMSD of each trajectory frame against the prepared X-ray crystal structure and RMSF from the average structure for each residue. NAP exhibited a similar RMSD with backbone of Rab5, suggesting that this compound forms stable complex with Rab5 (Fig. 3D). The RMSF plot showed the fluctuation of the residues compared to the average structure across the Rab5-NAP/GDP complex trajectory, showing that the two complexes have similar flexibility (Fig. 3E).

To investigate the interaction of NAP and Rab5 in detail, we analyzed the non-bond interactions between ligands (NAP, GDP and GTP) and receptor Rab5. Top 10 poses of the three ligands interacted with 18 residues of Rab5 through diverse bindings were compared according to their formed bond numbers. As shown in the heatmap of the hydrophobic and hydrogen-bond interactions, we found that NAP exhibited similar interaction patterns to GDP or GTP in residues Ala30, Lys33 and Ser34 of Rab5, but not in other regions of Rab5 (Fig. 4A). Since hydrogen interaction is the most important form of drug-protein interaction [20], we analyzed their hydrogen-bond counts and found that Ser34 formed the greatest number of hydrogen bonds with NAP and GDP/GTP (Fig. 4B-C), suggesting that NAP may competitively bind to Ser34 to disrupt GDP/GTP exchange in Rab5.

We chose the top docking pose of Rab5-NAP to show proposed interactions between NAP and Rab5 residues using 2D diagram (Fig. 4D). As such, we can see that residues Ala30, Val31, Gly32, Lys33 Ser34, Glu80, Asn133 and Asp136 could form hydrogen bonds or Pi-Alkyl with NAP. And we then *in silico* mutated these

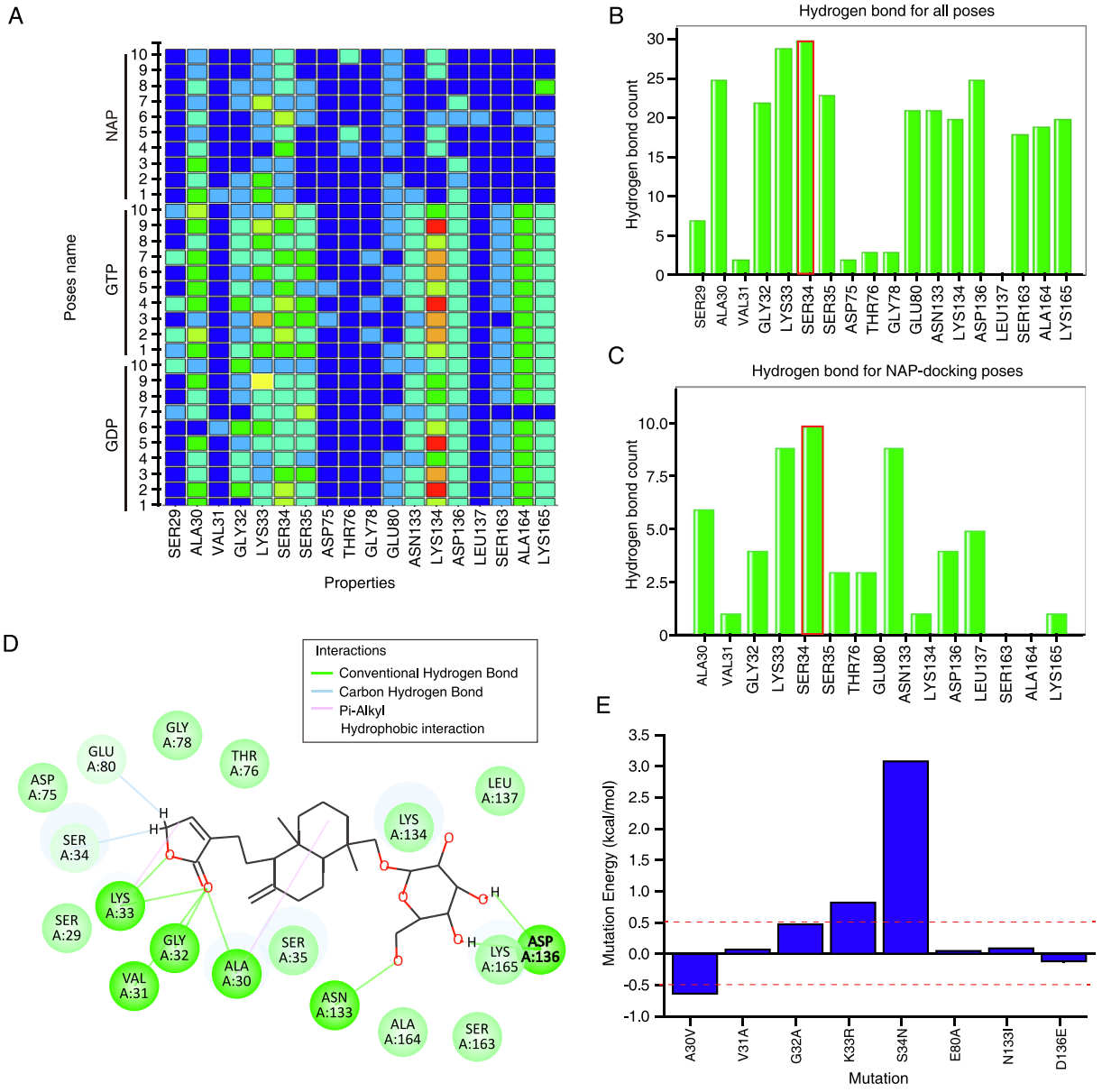


Fig. 4. Details of the possible interactions between NAP and Rab5. (A) Heatmap summarizing the interaction counts and listing the most frequently encountered residues for each interaction type. (B, C) Frequency histograms summarizing the hydrogen-bond interactions calculated for all the docked poses (B) and the NAP-docked poses (C). (D) 2D diagram showing proposed interactions between NAP and Rab5 residues: Conventional hydrogen bonds were represented as green lines, carbon hydrogen bonds were represented as blue lines and pi-alkylare bonds were represented as pink lines. (E) Mutation energy for the Rab5 residues that could form hydrogen bonds with NAP. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

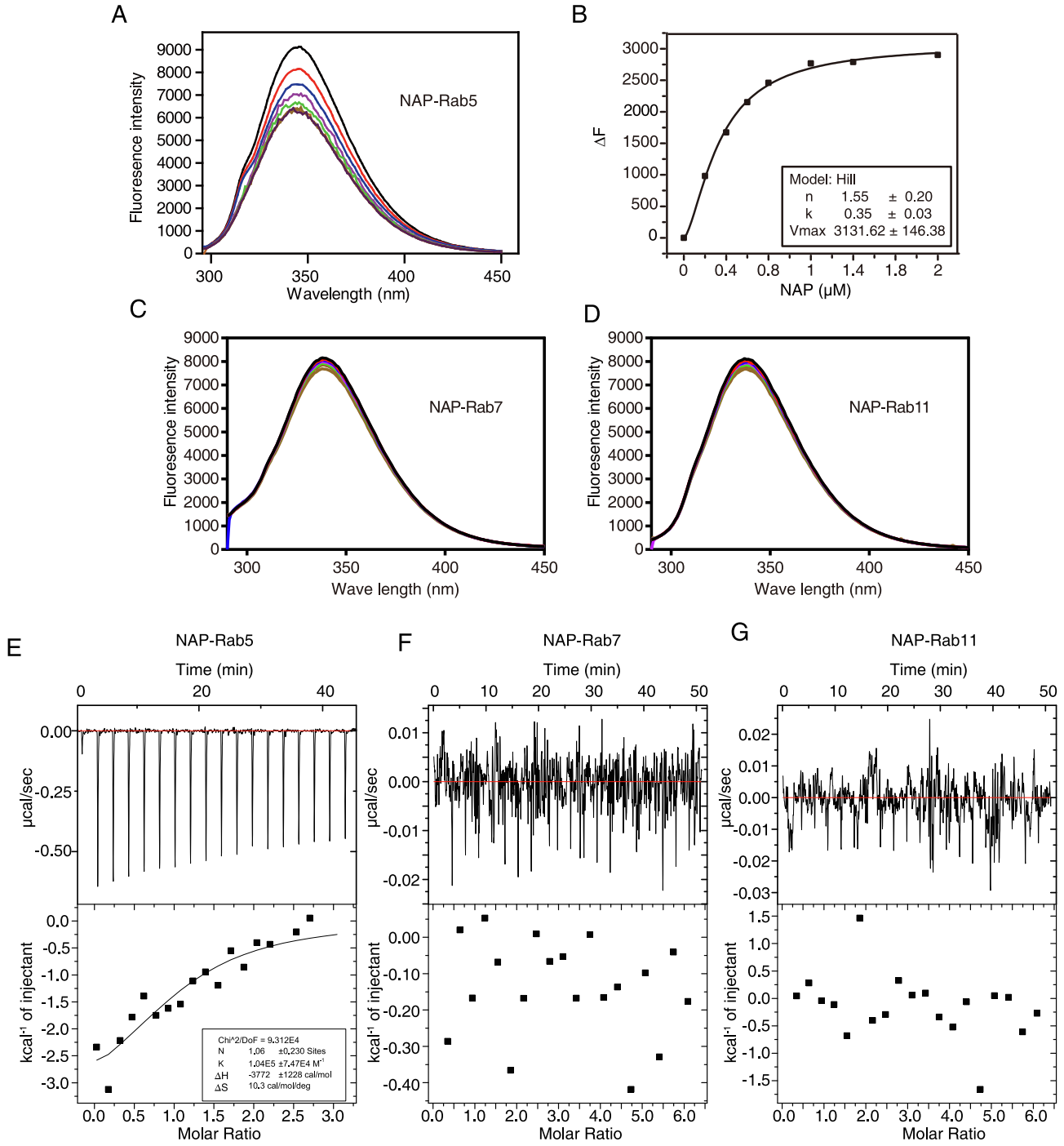


Fig. 5. NAP binds to Rab5 only. (A) Fluorescence titration experiment showing a strong interaction between NAP and Rab5. (B) The fluorescence quenching at 330 nm *verses* NAP concentration was fitted with Hill plot. (C, D) Fluorescence titration experiment showing no interaction between NAP and Rab7 (C) or Rab11 (D). (E-G) Isothermal titration calorimetry binding curves of NAP-Rab5 (E), NAP-Rab7 (F) and NAP-Rab11 (G). The parameters N, KD, ΔH, and ΔS are shown in the diagrams.

amino acids individually to investigate their effects on Rab5 stability by calculating their mutation energy. As shown in Fig. 4E, mutation of Ala30 could enhance Rab5 stability with mutation energy less than −0.5 kcal/mol, mutating Gly32, Lys33 and Ser34 could significantly destabilize Rab5 structure with mutation energy greater than 0.5 kcal/mol, while mutations at other residues had little impact (Fig. 4E). Especially, Ser34 mutation showed the most dramatic increase of mutation energy, suggesting that Ser34 is the most important residue for Rab5-ligand binding stability.

3.4. NAP inhibits Rab5 activity

To support the direct physiological interaction between NAP and Rab5, we isolated pure Rab5 protein for *in vitro* fluorescence titration experiment. As shown in Fig. 5A-B, titration of NAP to the purified Rab5 solution showed a stepwise fluorescence quenching at around 340 nm, with an equilibrium dissociation constant Kd = 0.35 ± 0.03 μM. Interestingly, no significant fluorescence change was found in Rab7 or Rab11 titrated with NAP (Fig. 5-

C-D), suggesting a specific binding of NAP with Rab5. We then employed isothermal titration calorimetry (ITC) to further verify this direct interaction and found that NAP significantly bound to Rab5, but not Rab7 or Rab11 (Fig. 5E-G). We next performed GST-R5BD pulldown assay to assess whether the NAP binding influences the activity of Rab5. The results in Fig. 6A revealed that NAP decreased the activity of Rab5 in a dose-dependent manner (up to 150 μM), where EGF treatment was set as a positive control, suggesting that the binding of NAP to Rab5 may disrupt its GTP/GDP exchange, diminishing the active form of Rab5, and subsequently p-ERK.

We then investigated the effect of NAP on Rab5 function in cells. Rab5 is the key regulator in endosome trafficking. Upon activation by its guanine nucleotide exchange factor Rin1, Rab5 participates EGFR internalization, degradation and signal transduction to promote cell growth [21]. Our experiment shown in Fig. 6B demonstrated that EGFR degradation was delayed when cells were treated with NAP. In addition, FITC-Dextran uptake assay showed that NAP could induce the accumulation of FITC signal in A549 in dose- and time-dependent manners (Fig. 6C), suggesting an inhibitory effect of NAP on Rab5 activation, leading to the handicap of EGFR internalization and uptake ability. We further detected the effect of NAP on cancer cell growth by using WST-1 and colony formation assays. As shown in Fig. 6D, A549 cells were treated with increasing concentrations of NAP (up to 400 μM) for 48 h, the cell viability was significantly diminished. The effect of long-term treatment of NAP (up to 100 μM) in A549 for 15 days was detected by colony formation assay (Fig. 6E), showing that NAP inhibited the growth of A549 cells in a dose-dependent manner.

As discussed above, Ser34 is a critical amino acid in Rab5 that has the most important interaction with the substrates, we then tested the effect of the mutation at Ser34 on the function of Rab5. As expected, the mutation from serine to asparagine (S34N) abolished the GTP binding with Rab5, resulting in the accu-

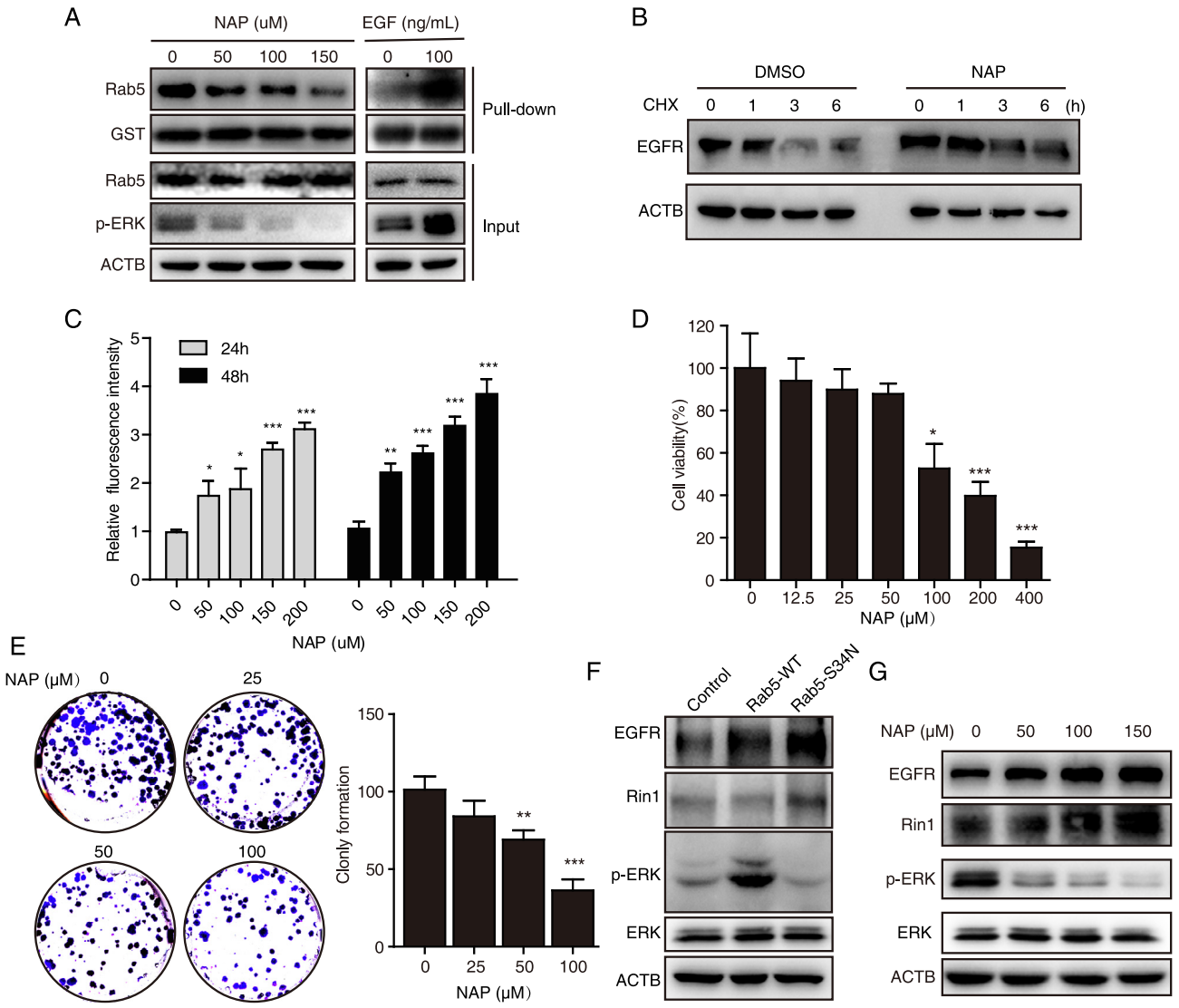


Fig. 6. NAP suppresses the activity of Rab5 and the subsequent phenotypes (A) A549 cells were treated with indicated concentration of NAP, and the Rab5 activity was detected by using GST-R5BD pull-down assay. (B) A549 cells were pretreated with NAP or DMSO for 48 h, and followed by cycloheximide (CHX, 50 μg/mL) treatment for 1, 3, and 6 h. Cells were harvested at the indicated time points and EGFR levels were determined by Western blotting. (C) A549 cells were loaded with FITC-Dextran for 48 h and then treated with increasing concentration of NAP for another indicated hours, the fluorescence of FITC was detected by flow cytometry. (D) A549 cells were incubated with indicated concentrations of NAP for 48 h, the cell viability was determined by WST-1 assays. (E) Colony formation assays were performed to determine the long-term treatment of NAP in A549 cells. (F) A549 cells were transfected with the plasmid expressing Rab5-WT or Rab5-S34N as indicated, and expression levels of EGFR, Rin1 and p-ERK were detected by Western blotting. (G) A549 cells were treated with indicated concentrations of NAP, and expression levels of EGFR, Rin1 and p-ERK were detected. Bars, SD; N = 3; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

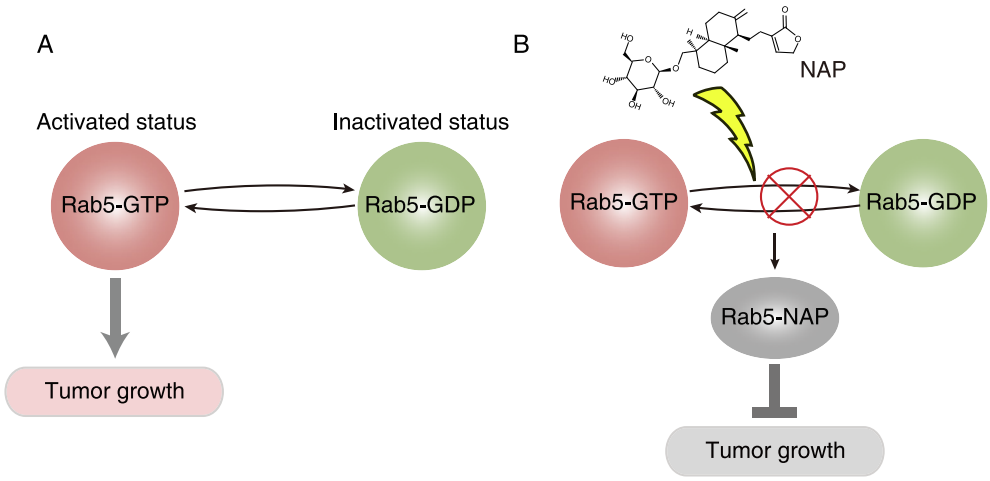


Fig. 7. Schematic diagram of NAP on disrupting GDP/GTP exchange and the activity of Rab5. (A) Rab5 exerts its biological function by switching GDP- and GTP-bound states, where Rab5-GTP promotes tumorigenesis. (B) NAP disrupts GDP/GTP exchange in Rab5, exhibiting anticancer effect.

mulation of EGFR and Rin1, and the decrease of p-ERK (Fig. 6F). Accordingly, NAP treatment in A549 cell achieved similar effects on the accumulation of EGFR and Rin1 and the decrease of p-ERK (Fig. 6G).

4. Conclusions

In summary, we have established an efficient strategy to screen for Rab5 inhibitors that can specifically occupy its surface GDP/GTP binding groove using natural product library, and identified NAP as a novel Rab5 inhibitor with optimal ADMET property. NAP directly binds to Rab5 *via* forming hydrogen bonds with Ala30, Val31, Gly32, Lys33, Ser34, Glu80, Asn133 and Asp136 in the binding pocket on Rab5 surface, disrupting the GDP/GTP exchange and thus decreasing the activity of Rab5 to inhibit the cancer cell growth (Fig. 7). This finding breaks the miserable situation on undruggable Rab5 and highlights the potential of NAP for further development as a chemotherapeutic agent for cancer therapy.

Author statement

Y.W., J.Z. and Q.Y.H. designed the experiments and analyzed results. J.Z. carried out the experiments and wrote the manuscript. J.Z., Y.S. and L.Y.Z. assisted in the protein purification, ITC and fluorescence titration experiments. N.N.Y. and L.O.Y. assisted in GST pull down and Western blotting assay. J.Z. assisted with the analysis of computational experiments. R.D.F. assisted with the analysis of data; Y.W. and Q.Y.H. supervised the research and edited the manuscript. All authors discussed the results and commented on the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Key R&D Program of China (2017YFA0505100), China Postdoctoral Science Foundation (2018M643372, 2020T130252), The National Natural Science Foundation of China (82002948, 31770888), Guangdong Natural Science Research Grant (2019A1515010196, 2019A1515110597), Guangdong "Climbing Program" Special Funds (pdjh2020a0064) and the Fundamental Research Funds for the Central Universities (11619303).

References

[1] Stenmark H. Rab GTPases as coordinators of vesicle traffic. Nat Rev Mol Cell Biol 2009;10(8):513–25. https://doi.org/10.1038/nrm2728.
[2] Barbieri MA, Fernandez-Pol S, Hunker C, Horazdovsky BH, Stahl PD. Role of Rab5 in EGF receptor-mediated signal transduction. Eur J Cell Biol 2004;83(6):305–14. https://doi.org/10.1078/0171-9335-00381.
[3] Chen P-I, Kong C, Su X, Stahl PD. Rab5 isoforms differentially regulate the trafficking and degradation of epidermal growth factor receptors. J. Biol. Chem. 2009;284(44):30328–38. https://doi.org/10.1074/jbc.M109.034546.
[4] Cong XX, Gao XK, Rao XS, Wen J, Liu XC, Shi YP, He MY, Shen WL, Shen Y, Ouyang H, Hu P, Low BC, Meng ZX, Ke YH, Zheng MZ, Lu LR, Liang YH, Zheng LL, Zhou YT. Rab5a activates IRS1 to coordinate IGF-AKT-mTOR signaling and myoblast differentiation during muscle regeneration. Cell Death Differ 2020;27(8):2344–62. https://doi.org/10.1038/s41418-020-0508-1.
[5] Mendoza P, Ortiz R, Diaz J, Quest AFG, Leyton L, Stupack D, Torres VA. Rab5 activation promotes focal adhesion disassembly, migration and invasiveness in tumor cells. J Cell Sci 2013;126(17):3835–47. https://doi.org/10.1242/jcs.119727.
[6] Igarashi T, Araki K, Yokobori T, Altan B, Yamanaka T, Ishii N, Tsukagoshi M, Watanabe A, Kubo N, Handa T, Hosouchi Y, Nishiyama M, Oyama T, Shirabe K, Kuwano H. Association of RAB5 overexpression in pancreatic cancer with cancer progression and poor prognosis via E-cadherin suppression. Oncotarget 2017;8(7):12290–300. https://doi.org/10.18632/oncotarget.14703.
[7] Yang PS, Yin PH, Tseng LM, Yang CH, Hsu CY, Lee MY, et al. Rab5A is associated with axillary lymph node metastasis in breast cancer patients. Cancer science. 2011;102:2172-8.
[8] Zarei O, Sarri N, Dastmalchi S, Zokai F, Papadopoulos N, Lennartsson J, Heldin C-H, Hamzeh-Mivehroud M. Structure-based discovery of novel small molecule inhibitors of platelet-derived growth factor-B. Bioorg Chem 2020;94:103374. https://doi.org/10.1016/j.bioorg.2019.103374.
[9] Wang L, Li Li, Zhou Z-H, Jiang Z-Y, You Q-D, Xu X-L. Structure-based virtual screening and optimization of modulators targeting Hsp90-Cdc37 interaction. Eur J Med Chem 2017;136:63–73. https://doi.org/10.1016/j.ejmech.2017.04.074.
[10] Fatima S, Gupta P, Sharma S, Sharma A, Agarwal SM. ADMET profiling of geographically diverse phytochemical using chemoinformatic tools. Future Med Chem 2020;12(1):69–87. https://doi.org/10.4155/fmc-2019-0206.
[11] Wang Y, Yu RY, He QY. Proteomic analysis of anticancer TCMs targeted at mitochondria. Evid Based Complement Alternat Med. 2015;2015:539260.
[12] Wang Y, Zhang J, Huang ZH, Huang XH, Zheng WB, Yin XF, et al. Isodeoxyelephantopin induces protective autophagy in lung cancer cells via Nrf2-p62-keap1 feedback loop. Cell Death Dis 2017;8:e2876.
[13] Cai F, Chen M, Zha D, Zhang P, Zhang X, Cao N, et al. Curcumol potentiates celecoxib-induced growth inhibition and apoptosis in human non-small cell lung cancer. Oncotarget 2017;8:115526–45.
[14] Wang Y, Li YJ, Huang XH, Zheng CC, Yin XF, Li B, et al. Liensinine perchlorate inhibits colorectal cancer tumorigenesis by inducing mitochondrial dysfunction and apoptosis. Food Funct 2018;9:5536–46.
[15] Zhang L, Li N, Cao K, Yang XY, Zeng G, Sun X, et al. Crucial residue Trp158 of lipoprotein PiaA stabilizes the ferrichrome-PiaA complex in Streptococcus pneumoniae. J Inorg Biochem 2017;167:150–6.
[16] Zhang J, Zhou Y, Li N, Liu WT, Liang JZ, Sun Y, et al. Curcumol overcomes TRAIL resistance of non-small cell lung cancer by targeting NRH: quinone oxidoreductase 2 (NQO2). Adv Sci 2020.

[17] Zhang J, Wang Y, Zhou Y, He QY. Jolkinolide B induces apoptosis of colorectal carcinoma through ROS-ER stress-Ca(2+)-mitochondria dependent pathway. Oncotarget. 2017;8:91223–37.
[18] Yuan W, Song C. The emerging role of Rab5 in membrane receptor trafficking and signaling pathways. Biochem Res Int 2020;2020:4186308.
[19] Yang JF, Wang F, Chen YZ, Hao GF, Yang GF. Erratum to: LARMD: integration of bioinformatic resources to profile ligand-driven protein dynamics with a case on the activation of estrogen receptor. Briefings Bioinf 2020.
[20] Chen D, Oezguen N, Urvil P, Ferguson C, Dann SM, Savidge TC. Regulation of protein-ligand binding affinity by hydrogen bond pairing. Sci Adv 2016;2:e1501240.
[21] Barbieri MA, Roberts RL, Gumusboga A, Highfield H, Alvarez-Dominguez C, Wells A, et al. Epidermal growth factor and membrane trafficking. EGF receptor activation of endocytosis requires Rab5a. J Cell Biol 2000;151:539–50.