analytical chemistry

pubs.acs.org/ac Article

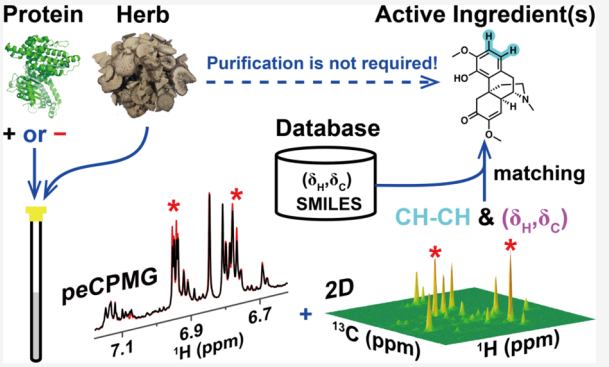
Rapid Targeted Screening and Identification of Active Ingredients in Herbal Extracts through Ligand-Detected NMR and Database Matching

Tao Huang, Xin Chai, Shuangli Li, Biao Liu, Jianhua Zhan, Xiaohua Wang, Xiong Xiao, Qinjun Zhu, Caixiang Liu, Danyun Zeng, Bin Jiang, Xin Zhou, Lichun He, Zhou Gong, Maili Liu,* and Xu Zhang*

Cite This: *Anal. Chem.* 2024, 96, 15194−15204 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Herbal extracts are rich sources of active compounds that can be used for drug screening due to their diverse and unique chemical structures. However, traditional methods for screening these compounds are notably laborious and time-consuming. In this manuscript, we introduce a new high-throughput approach that combines nuclear magnetic resonance (NMR) spectroscopy with a tailored database and algorithm to rapidly identify bioactive components in herbal extracts. This method distinguishes characteristic signals and structural motifs of active constituents in the raw extracts through a relaxation-weighted technique, particularly utilizing the perfect echo Carr−Purcell−Meiboom−Gill (peCPMG) sequence, complemented by precise 2D spectroscopic strategies. The cornerstone of our approach is a customized database designed to filter potential compounds based on defined parameters, such as the presence of CH_n segments and unique chemical shifts, thereby expediting the identification of promising compounds. This innovative technique was applied to identifying substances interacting with choline kinase α (ChoKα1), resulting in the discovery of four new inhibitors. Our findings demonstrate a powerful tool for unraveling the complex chemical landscape of herbal extracts, considerably facilitating the search for new pharmaceutical candidates. This approach offers an efficient alternative to traditional methods in the quest for drug discovery from natural sources.



INTRODUCTION

Traditional herbal medicine has been used for thousands of years and remains popular worldwide due to its effectiveness in treating various diseases.[1] Medicinal plants contain abundant phytochemicals that can be used to source lead compounds in drug discovery, as they have a wide range of biological activities.[2] In fact, over the last four decades, over half of all approved small-molecule drugs have either been derived from natural products or connected to them.[3] This demonstrates the important role that plant-based resources play in producing new medicinal agents and highlights their continued importance in pharmaceutical development.

Lead discovery from herbal sources primarily involves two strategies: the traditional activity-guided approach and the modern ligand-fishing method. The traditional approach has extensive applications but is often inefficient due to the laborious extraction, purification, and structural identification processes. On the other hand, the ligand-fishing approach leverages the specific binding affinities between target proteins and potential ligands.[4] To streamline the isolation and characterization of active constituents, this strategy utilizes various techniques, such as magnetic bead capture,[5−7] affinity ultrafiltration,[8−10] affinity chromatography,[11,12] molecular exclusion,[13] and affinity capillary electrophoresis.[13] However, limitations arise when considering weak-binding components and the need for reference standards, which often necessitates extensive compound isolation for structural identification. Consequently, developing more efficient methods to screen and identify active compounds in herbal sources is an ongoing pursuit.

Nuclear magnetic resonance (NMR) has risen as a pivotal tool in drug discovery and molecular structure analysis.[14,15] Its utility in drug discovery is well-recognized, with techniques such as saturation-transfer difference (STD) and water−ligand observation with gradient spectroscopy (WaterLOGSY) facilitating the investigation of molecular interactions[16−19] and analysis of complex mixtures.[20,21] It has shown remarkable promise, as demonstrated by studies that have successfully

Received: April 30, 2024
Revised: August 22, 2024
Accepted: August 29, 2024
Published: September 12, 2024

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

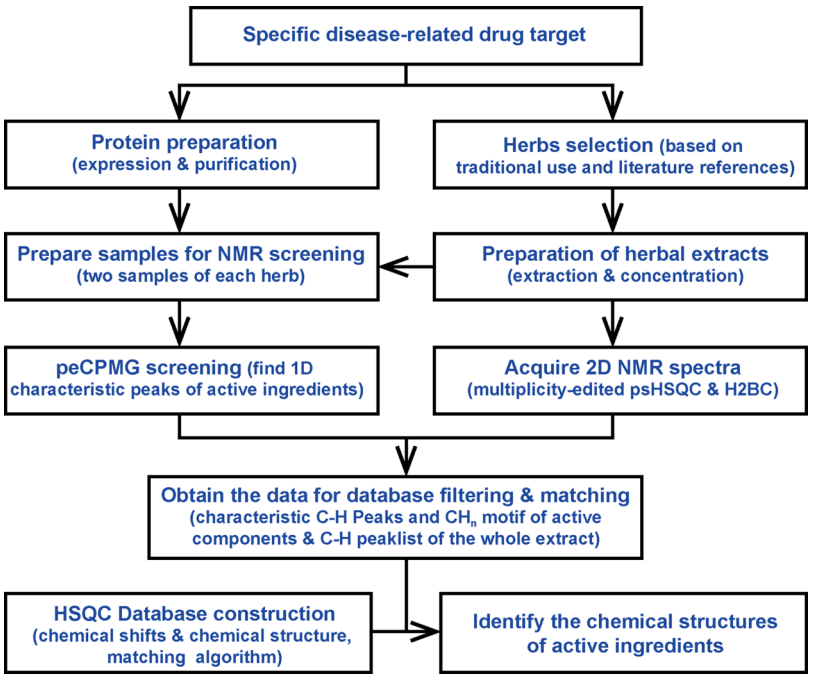


Figure 1. Flowchart depicting the rapid screening and identification process for active ingredients in herbal medicines.

pinpointed bioactive compounds in herbal medicine using this method.[22−25]

Despite these advancements, structural elucidation still poses challenges, often requiring advanced separation or specialized expertise. Nevertheless, a rich repository of plant-derived natural product structures and their associated NMR data provides a solid foundation for reference comparison,[26] aided by databases dedicated to natural product NMR information, such as NAPROC-13,[27] NMRShiftDB2,[28] JEOL CH-NMR-NP,[29] and NP-MRD.[26] Moreover, the development of automated methods for metabolite identification in metabolomics, such as COLMAR,[20] MetaboMiner,[30] PROMEB,[31] and NPid,[21] gives NMR great potential to identify molecules in herbals quickly.

In this study, we present a rapid method for targeted screening and identification of bioactive components in crude herbal extracts by combining ligand-detected NMR with database matching, as illustrated in Figure 1. The sample preparation is similar to STD or waterLOGSY NMR experiments. However, identification of the active components within the herbal extracts is facilitated by a novel relaxation-weighted spectroscopy accessed by using a perfect echo Carr−Purcell−Meiboom−Gill (peCPMG) sequence, which is depicted in Figure S1. Additional insights into the active components, including characteristic C−H peaks and CH_n structural motifs, are obtained using 2D multiplicity-edited pure shift Heteronuclear Single Quantum Coherence (edited psHSQC) and Heteronuclear 2-Bond Correlation (H2BC) spectra. This information is then processed through a specialized NMR database that we developed, enabling the filtering of potential compounds based on criteria including the presence of CH_n motifs in their chemical structure and the alignment of all characteristic C−H peaks in terms of chemical shifts. Potential active agents in the complex mixture are further narrowed down by calculating a score based on the ratios of matched peaks to the total number of peaks. The validation and efficacy of this novel method have been successfully demonstrated in the screening for active compounds targeting choline kinase α (ChoKα1), including known inhibitors, leading to the discovery of four inhibitors with novel scaffolds. Our approach has proven effective in identifying active components with varying binding affinities to ChoKα1 across a wide concentration range, offering a potent tool for the rapid screening and identification of bioactive components in herbal medicine.

MATERIALS AND METHODS

Database Design and Implementation. Our statistical analysis relied on natural product structures obtained from the COCONUT database.[32] We initially screened 407,270 compounds, applying filters to include only structures containing the elements C, H, O, and N, with a maximum logP value of 5. This process resulted in a data set of 261,278 structures eligible for the subsequent CH_n motif analysis.

For representation purposes, chemical structures were coded as SMILES (Simplified Molecular Input Line Entry System) strings,[33,34] with CH_n motifs identified using the SMARTS (SMiles ARbitrary Target Specification) language.[35] The identification and extraction of the CH_n motifs were performed using the RDKit toolkit, available at www.rdkit.org, within the Python programming environment, version v3.10, accessible at www.python.org.

The NMR data used in this study were primarily sourced from the CH-NMR-NP database and supplemented by data gleaned from published scientific articles.[29] It is worth noting that all NMR data utilized here have been previously disclosed in peer-reviewed scientific literature.

We harnessed ^{1}H and ^{13}C chemical shifts, along with their respective assignments, to compile HSQC peak tables. Additionally, each HSQC peak table features a separate CH_n column, denoting the C−H coupling, where "n" represents the number of hydrogen atoms attached to the carbon atom in that group.

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

Our database entries contain a wealth of information, including an HSQC peak table, the compounds' names, their CAS registry number, chemical structure, the NMR solvent used, and bibliographical references (further detailed in the Supporting Information). Currently, the database encompasses 23,660 natural products.

Protein Expression and Purification. To express human Chokα1, a construct was utilized that included an N-terminal His-tag, a flexible G_4S linker, and a TEV protease cleavage site featuring the sequence M-HHHHHH-GGGGS-ENLYFQG-hChoK$\alpha 1_{75-457}$. Expression in *Escherichia coli* was initiated by incubating the cells at 37 °C until an $OD_{600\,nm}$ of 0.8 was reached. Induction of protein expression was then triggered with 0.5 mM IPTG (isopropyl β-D-1-thiogalactopyranoside), after which incubated continued overnight at 20 °C. Harvesting of the cells occurred via centrifugation at 6000 rpm for 10 min, followed by resuspension in buffer A (20 mM phosphate/50 mM NaCl, pH 7.4). Cells disruption was through sonication, and subsequent centrifuged at 20,000 rpm at 4 °C for 30 min helped to isolate supernatants. These were applied to a HisTrap HP column (GE Healthcare). Washed steps involved buffer B (20 mM phosphate/50 mM NaCl/50 mM imidazole, pH 7.4), with Chokα1 elution in buffer C (20 mM phosphate/50 mM NaCl/400 mM imidazole, pH 7.4). Merged fractions containing the enzyme were concentrated, and buffer exchanged into buffer A using a 30,000 kDa MWCO concentrator (Millipore). SDS-PAGE and mass spectrometry analyses confirmed protein purity and identity. Enzymatic activity was verified via phosphorylation of choline and protein concentration determined through absorbance at 280 nm with an extinction coefficient of 60,195 M^{-1} cm^{-1}.

Sample Preparation. Herbal materials were purchased from a specialized Chinese medicinal e-commerce platform (https://www.zyccst.com/). Herbs were pulverized and then soaked in methanol overnight. Soaking included sonication for the initial and final 10 min. Afterward, the extract was filtered and concentrated under reduced pressure until dry. The total alkaloid extract of *Stephania tetrandra* was obtained through a combination of acid−base treatment and solvent extraction. For NMR analysis, 55 mg of dry extract dissolved in 550 μL DMSO-d_6 underwent vortexing and sonication in a water bath for two 5 min cycles. Centrifugation at 15,000 rpm for 5 min followed, with the supernatant serving as the stock solution for further tests.

Active Ingredient Screening. A sample for active ingredient screening was prepared by combining 10 μL of the herbal stock solution with 10 μL DMSO-d_6, 50 μL D_2O, 478 μL of a 50 mM phosphate buffer (pH 7.4), and 2 μL of a 30 mM TSP solution. After a 30-s vortex and 5 min centrifugation at 15,000 rpm, the supernatant was mixed with either Chokα1 (10 μM, final concentration) or an equivalent amount of phosphate buffer. The samples were then transferred to a 5 mm NMR tube for subsequent analysis.

Enzymatic Activity Assays. To access the inhibitory effects of the screened active components, in vitro enzyme activity experiments were carried out. Chokα1 was mixed into a 50 mM Tris-HCl buffer (pH 7.4) containing 1.8 mM $MgCl_2$, 5.4 mM ATP, 3.6 mM choline, and 50 μL D_2O, with a final enzyme concentration of 1 μM. Separate samples contained either 20 μL of DMSO-d_6 control or the test compound stock, with a final concentration of 360 μM. The reaction mixture was incubated at room temperature for 8 min before a thermal denaturation halt step in a boiling water bath for 5 min. Afterward, an NMR tube housed the supernatant for choline and phosphocholine concentration comparisons using the 1D NOESY pulse sequence with presaturation to determine any inhibitory effect.

NMR Spectroscopy. All NMR spectra were acquired at 25 °C on a 600 MHz Bruker Avance III HD spectrometer equipped with a 5 mm TCI CryoProbe (Bruker Biospin, Germany). Ligand screening employed the T_2-weighted Carr−Purcell−Meiboom−Gill (CPMG) sequence due to the shortening of ligand T_2 relaxation times upon binding to the target. To mitigate phase and multiplet distortions,[36] as shown in Figure S2, a "perfect echo" scheme was applied in conjunction with CPMG and Watergate-W5 modules[37,38] (Figure S1). Both gradients g1 and g2 were smoothed, square-shaped, 1 ms long, with a subsequent 200 μs recovery delay. The total CPMG time was approximately 100 ms (τ = 400 μs, n = 64). The ^{1}H carrier frequency was aligned with the water resonance, and the spectral width was 20 ppm. A total of 32k complex points were collected, and 32 scans were accumulated with a recycle delay of 2 s. FID processing included a 0.3 Hz exponential broadening, "qfil" baseline correction across 0.2 ppm, zero-filled to 64k complex points, and manual phase and baseline correction. Chemical shifts referenced the TSP signal at 0 ppm.

To effectively differentiate various CH_n signals, we employed a multiplicity-edited pure shift HSQC sequence (edited psHSQC), incorporating BIRD (Bilinear Rotation Decoupling) for refocusing active spin within the HSQC experiments.[39] The multiplicity editing allowed us to effectively discern CH_2 signals, which manifested as negative peaks, from the positive peaks corresponding to CH and CH_3 signals. Furthermore, the sequence facilitated the conversion of multiplet resonances into singlets by eliminating the impact of homonuclear couplings, except for nonequivalent CH_2 signals, which retained their doublet appearance. This streamlining enhanced our ability to accurately measure peak intensities. The experimental parameters for this edited psHSQC were consistent with our previously described psHSQC methodology.[21] After spectral acquisition, we employed the NMRFAM-SPARKY software for peak picking and deconvolution with pseudo-Voigt line profiles. To elucidate the adjacency of C−H groups, we conducted H2BC experiments as described in earlier publications.[21,40]

For STD experiments, we employed a pseudo-2D version of the STD NMR sequence, which was equipped with a perfect double-W5-echo scheme for efficacious water suppression during the interleaved acquisition of on- and off-resonance spectra. The relaxation delay was set to 4 s, during which the saturation pulses were applied specifically targeting the protein, via a train of 50 ms Gaussian pulses at a 200 Hz bandwidth to ensure comprehensive saturation. We included a 25 ms spinlock pulse at a 10 kHz field strength before acquisition to remove residual protein resonances. A homospoil sequence was introduced after each scan to destroy any unwanted magnetization. Saturation times spanned multiple durations (0.5, 0.75, 1.0, 1.5, 2.0, and 3.0 s) to facilitate the generation of an STD buildup curve conducive to fitting dissociation constants (K_D). In general, 64 scans were recorded, with increases in powers of 2 implemented for cases featuring short saturation time or diminished ligand concentrations. The initial rate model fitting to obtain the K_D was based on peak height measurements.[41]

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

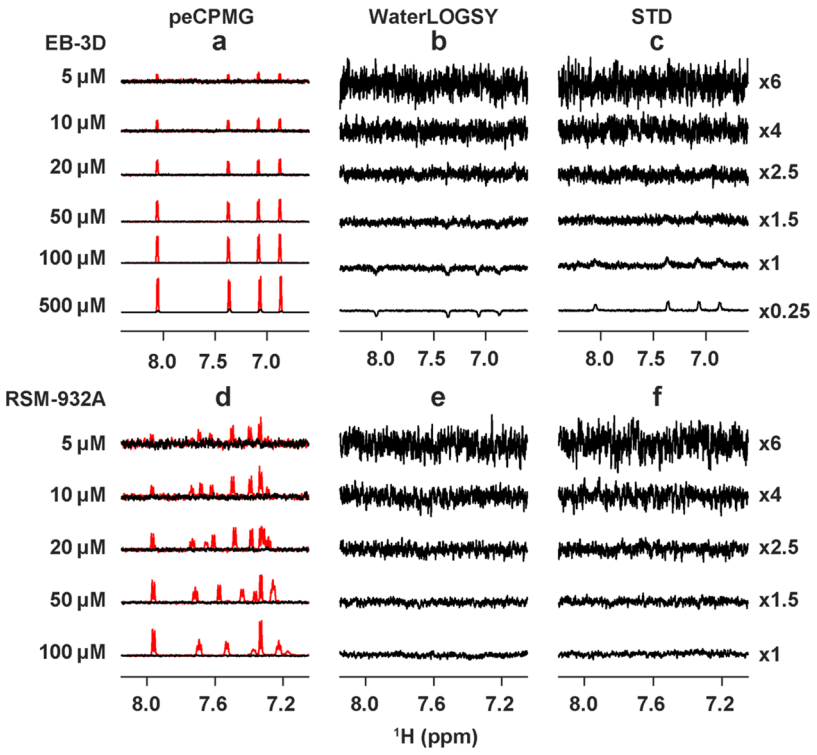


Figure 2. Comparative spectra of EB-3D (panels a, b, and c) and RSM-932A (panels d, e, and f) using peCPMG, WaterLOGSY, and STD techniques with (black) and without (red) addition of 10 μM ChoKα1. All the spectra were acquired with 32 scans. Spectra normalization includes a scaling factor detailed within the figure for enhanced visibility. Ligand concentrations are represented as follows: 5, 10, 20, 50, 100, and 500 μM. Note: Saturation of RSM-932A is observed at the concentration of 500 μM.

^{19}F-detected NMR competition experiments (FAXS) were conducted using single pulse excitation and inverse-gated ^{1}H decoupling. The ^{19}F and ^{1}H carrier frequencies were set to −120 and 4 ppm, respectively, and the spectral width was 200 ppm. A total of 64k complex points were collected, and 1024 scans were accumulated with a recycle delay of 1 s. The NMR samples contained 20 μM ChoKα1 in phosphate buffer, with 40 μM of spy molecules. Compounds were added at 400 μM final concentration from DMSO-d_6 stock.

Identification of the Active Components by Database Matching. Using peCPMG screening, we discern distinct ^{1}H signals that correspond to active components within the herbal extracts. These signals facilitate the precise determination of their specific C−H connectivities and the CH_n structural motifs through analysis of edited psHSQC and H2BC spectra. The identification of active components involves a two-phase approach.

Filtering Phase: Structural Motif and Characteristic C−H Peak Discrimination. Our first phase entails filtering potential active component structures against database entries, guided by the CH_n motifs and characteristic C−H peaks as benchmarks. We extract CH_n motifs represented in SMARTS notation, employing substructure searches with the RDkit software package. Candidate compounds must not only match CH_n structural motifs within their molecular architecture but must also display the full set of active component characteristic C−H peaks in their peak tables. Factoring variability influenced by solvent and temperature, we apply generous tolerances for ^{1}H and ^{13}C chemical shifts during the peak matching process. Only compounds fulfilling all these criteria progress as candidates for the further scoring phase.

Scoring Phase: Chemical Shift Matching. In this phase, we cross-reference each peak in the HSQC peak table of a candidate compound, drawn from the customized database, against the edited psHSQC peaklist of the herbal extract under examination. A candidate's specific C−H peak is considered concordant if a peak of edited psHSQC peaklist exists in its predefined chemical shift tolerance window and the peak intensity meets or exceeds 75% of the average intensity of the active component's characteristic C−H peaks. We compute the compatibility score for each candidate as follows

$$\text{score} = (N_m^{CH_2} + N_m^{CH_{1,3}})/N_t^{CH_n}$$

where $N_m^{CH_2}$ is the number of matched CH_2 peaks for a candidate, $N_m^{CH_{1,3}}$ signifies the number of matched CH and CH_3 peaks, and $N_t^{CH_n}$ refers the total number of HSQC peaks for the candidate. This scoring metric echoes methodologies employed in systems like COLMAR and NPid. Candidate compounds with high scores are considered likely active components, meriting further validation.

RESULTS AND DISCUSSION

^{1}H peCPMG NMR Screening Method. The peCPMG screening method was accessed for its validity and effectiveness using two known ChoKα1 ligands: EB-3D,[42] and RSM-932A.[43] Two peCPMG experiments can distinguish active components within a complex mixture by monitoring the changes in intensities or chemical shifts of protons. This is achieved by conducting experiments in the presence and absence of the target protein, which induces the binding of ligands.

Our results reveal that the presence of ChoKα1 causes noticeable reductions in the peCPMG spectral signals of

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

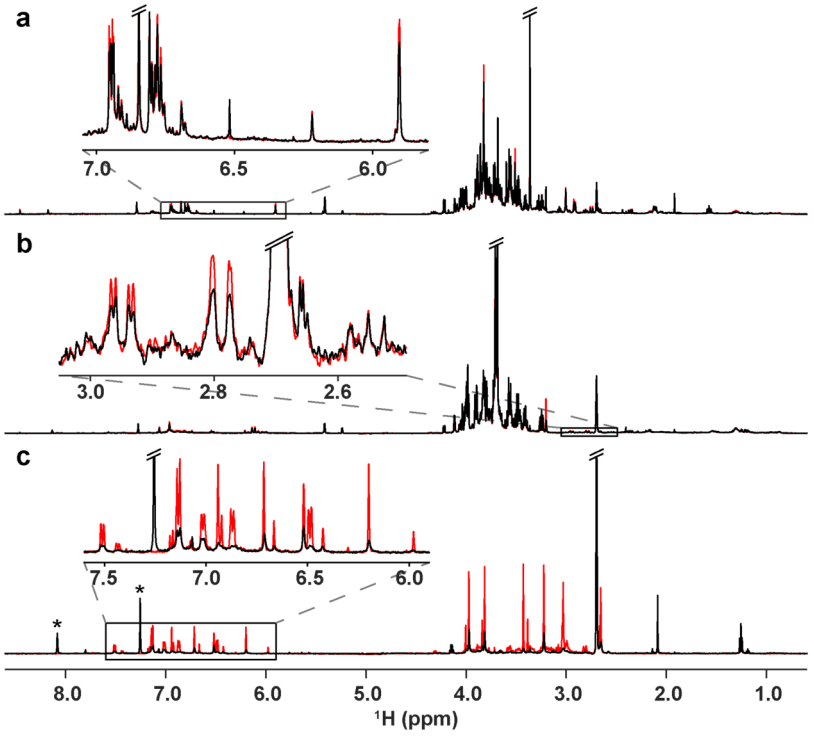


Figure 3. Overlayed ^{1}H peCPMG spectra of extracts in the presence (black) and absence (red) of ChoKα1. Target regions are emphasized with black rectangles and further magnifications are provided in inset images. ((a) methanol extract of *S. Caulis*. (b) methanol extract of *C. orbiculatus*. (c) Total alkaloid extract of *S. tetrandra*.). Note: Peaks indicated by an asterisk (*) represent residual imidazole signals in the protein solution.

ligands like EB-3D and RSM-932A. Such reductions facilitate the straightforward identification of active candidates (Figure 2). Furthermore, peCPMG displayed greater sensitivity in ligand-observed spectra compared to the commonly used STD and waterLOGSY methods, indicating its potential to identify candidates at low concentrations with varying affinities.

For example, the moderate-affinity ligand EB-3D can be screened using both WaterLOGSY and STD methods (see Figure 2b,c). However, these methods proved ineffective for high-affinity ligands like RSM-932A (Figure 2e,f). Nevertheless, the application of WaterLOGSY and STD is constrained by the necessity for an optimal ligand-to-target concentration ratio, even for low-affinity ligands. For EB-3D, interaction signals were undetectable when the ligand concentration was 5-fold lower than that of ChoKα1. (As shown in spectra of Figure 2b,c). The insufficiency of these NOE-based techniques is apparent for low-affinity, fragment-like compounds and is more pronounced for high-affinity ligands.

In contrast, peCPMG maintains higher sensitivity for both low- and high-affinity ligands, such as EB-3D and RSM-932A, compared to STD and WaterLOGSY under the same acquisition time, this suggests its versatility across a range of affinities (see Figure 2a,d). Given that herbal medicines consist of complex chemical constituents across a wide concentration range[44] (from about 10% down to 0.1%), our experimental setup encompassed approximately 1 mg of herbal extract in a volume of 550 μL, equivalent to concentration ranges of 5−500 μM. Irrespective of their binding affinity, any interacting components within herbal extract can be identified by scrutinizing the difference between peCPMG spectra obtained in the absence and presence of the target protein.

It is important to note that, for trace components constituting less than 0.1% in herbal extracts, one can implement conventional pretreatment methods to concentrate these components before NMR screening. For instance, enriching alkaloids, generally found at low concentrations, can be easily achieved through acid−base treatment combined with solvent extraction. Solid phase extraction (SPE) is another widely documented technique for enriching trace components in crude herbal extracts before NMR analysis.

Discovery of Active Components in Herbal Extracts through peCPMG NMR Screening. Our study aimed to demonstrate the practical efficacy of the proposed screening method in identifying new inhibitors from herbal extracts. ChoKα1, a viable therapeutic target, has been the focus of several inhibitors, notably HC-3 (hemicholinium-3) and RSM-932A, with the latter currently advancing through clinical trials.[45]

We explored three herbal extracts recognized for their inhibitory potential against ChoKα1 - *Sinomenii Caulis*, *Celastrus orbiculatus*, and *S. tetrandra*. Through the comparative analyses of ^{1}H peCPMG spectra of three extracts, both in the presence and absence of ChoKα1, we sought to pinpoint characteristic signals of active components (Figure 3). Upon integration with ChoKα1, a clear diminution in signal intensity was observed for one active compound in the T_2-weighted spectra of both *S. Caulis* and *C. orbiculatus* (Figure 3a,b). In contrast, the spectra for *S. tetrandra* displayed a broader array of active components, each exhibiting a decrease in peak intensity (Figure 3c). These modifications in NMR signals arise from an accelerated transverse relaxation effect, which occurs when active components bind to the target protein.

To further illustrate this phenomenon, we measured the transverse relaxation rates (R_2) of EB-3D and RSM-932A in

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

the absence and presence of the protein ChoKα1. Specifically, the methyl ^{1}H R_2 values of RSM-932A increased from 2.1 to 25 s^{-1} upon adding 10 μM ChoKα1 (Figure S10). Similarly, the R_2 values of methyl groups in EB-3D increased from 0.64 to 22 s^{-1} upon the addition of ChoKα1 (Figure S11). This binding event leads to the attenuation or disappearance of characteristic signals in the peCPMG spectrum, thus permitting the distinct identification of characteristic signals of active components within herbal extracts.

The ^{1}H peCPMG is a ligand-observed technique that leverages the transfer of protein-specific relaxation rates to bound small molecules. This method detects ligand-binding interactions through a decrease in signal intensity, which results from the transfer of the protein's short transverse relaxation time properties to the ligand. This effect is more pronounced with larger proteins, effectively eliminating theoretical upper size limitations.

Since STD and WaterLOGSY are primarily used for screening ligands with weak to moderate binding affinities, we conducted further experiments to demonstrate that peCPMG can also identify strongly binding components in herbal extracts. We utilized the extract of *S. Caulis* with the addition of the strong-binding ligand RSM-932A. Upon adding the protein to this extract, the peaks corresponding to RSM-932A disappeared (Figure S12). This result indicates that peCPMG can effectively screen for strongly binding components in herbal extracts.

It is important to note that peCPMG screening can encounter issues with nonspecific binding, such as the absorption of components to macromolecules. However, competition CPMG experiments can be utilized to avoid false positives resulting from nonspecific binding.[46] Additionally, complementary techniques such as FAXS,[47] STD epitope mapping,[48] and protein-observed fluorine (PrOF)[46] NMR methods can be employed to further exclude nonspecific binders and validate the specificity of the interactions.

A critical step in screening involves dissolving the components of herbal extracts within the protein solution. However, hydrophobic substances with limited water solubility may not effectively dissolve in aqueous media. The methanol extract of *S. Caulis* illustrates this challenge, with its lipid constituents remaining undetectable in phosphate buffer (Figure S3). Traditional medicinal herbal preparation methods, such as decoction, are known to enhance the solubility of active components.[49] Alternatively, DMSO has proven to be a solvent capable of enhancing solubility in aqueous buffers.[50] Given that most proteins tolerate concentrations of DMSO up to 10%, we propose increasing DMSO levels to address solubility issues of hydrophobic components found in herbal medicines, thereby improving the efficiency of our screening methodology.

Identification of Active Components' Structures in Herbal Extracts. To filter the candidate compounds, we use the characteristic C−H signals and CH_n motifs of each active component as the filtering criteria for database searches.

We investigated the role of CH_n motifs in the structural elucidation of active components within complex mixtures through a statistical analysis of natural product structures in the COCONUT database. Utilizing a custom RDkit script, we extracted CH_n motifs from each molecule in the COCONUT database. These motifs were combined and subjected to a substructure search within the database. We calculated hits-percentage by dividing the number of matched compounds by the total number of compounds in the database, with results depicted in Figure 4. The likelihood of encountering a motif

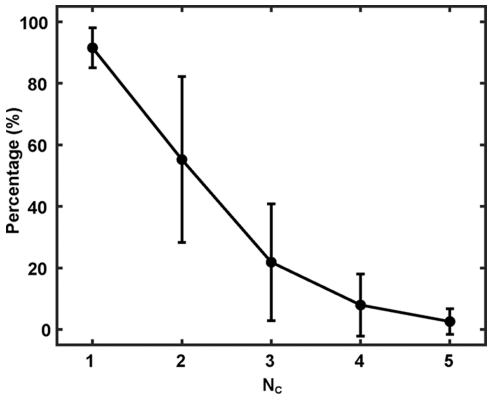


Figure 4. A statistical graph showing the relationship between query hits-percentage and the number of carbon atoms (N_C) in query CH_n motifs, along with error bars reflecting standard deviation across all corresponding hits-percentages. The analysis only includes CH_n motifs with up to 5 carbon atoms.

with only a single carbon atom is 92%. However, this probability drops to 19% for motifs containing three carbon atoms. This result indicates that as the number of carbon atoms in the CH_n motifs increases, the hits-percentage containing such motifs rapidly decreases, confirming that CH_n motifs are effective filtering criteria.

Edited psHSQC and H2BC experiments were conducted to align with database search criteria and algorithms, aiding in identifying corresponding C−H signals and CH_n motifs in the active components of herbal extracts. Structural identification must ascertain the origins of C−H peaks within the edited psHSQC spectrum, as they must originate from the same compound for accurate structural elucidation. After obtaining ^{1}H characteristic signals of the active component through ^{1}H peCPMG screening, the edited psHSQC spectrum was used to confirm the corresponding C−H signals. The H2BC spectrum then provided connectivity information for adjacent C−H groups, enabling us to identify structural CH_n motifs containing multiple contiguous C−H groups.

Applying these criteria, candidates from the database containing CH_n motifs and matching all C−H characteristic peaks in their HSQC peak table are retrievable. Additionally, a match score is computed for each candidate by comparing their HSQC peak table against the edited psHSQC peaklist of the herbal extract's components. This score indicates the likelihood of a candidate being the active component.

To exemplify this identification process, we examined the active component in the methanol extract of *S. Caulis*. Through ^{1}H peCPMG screening, three characteristic signals were noted (depicted in Figure 3a). These signals exhibited chemical shifts of 6.54, 5.70, and 6.74 ppm in the DMSO-d_6 solvent (indicated with black triangles in Figure S4). By correlating these ^{1}H signals, C−H signals in the edited psHSQC spectrum were assigned respectively (labeled with red asterisks in Figure S4). A direct linkage between the two C−H groups with ppm values of (6.74, 110.0) and (6.54, 118.0) was confirmed through H2BC correlations, validating the presence of a CHCH motif in the active compound. This directed the search to compounds with CHCH motif, resulting in 5911 candidates. Further refining for exact chemical shifts matches reduced this

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

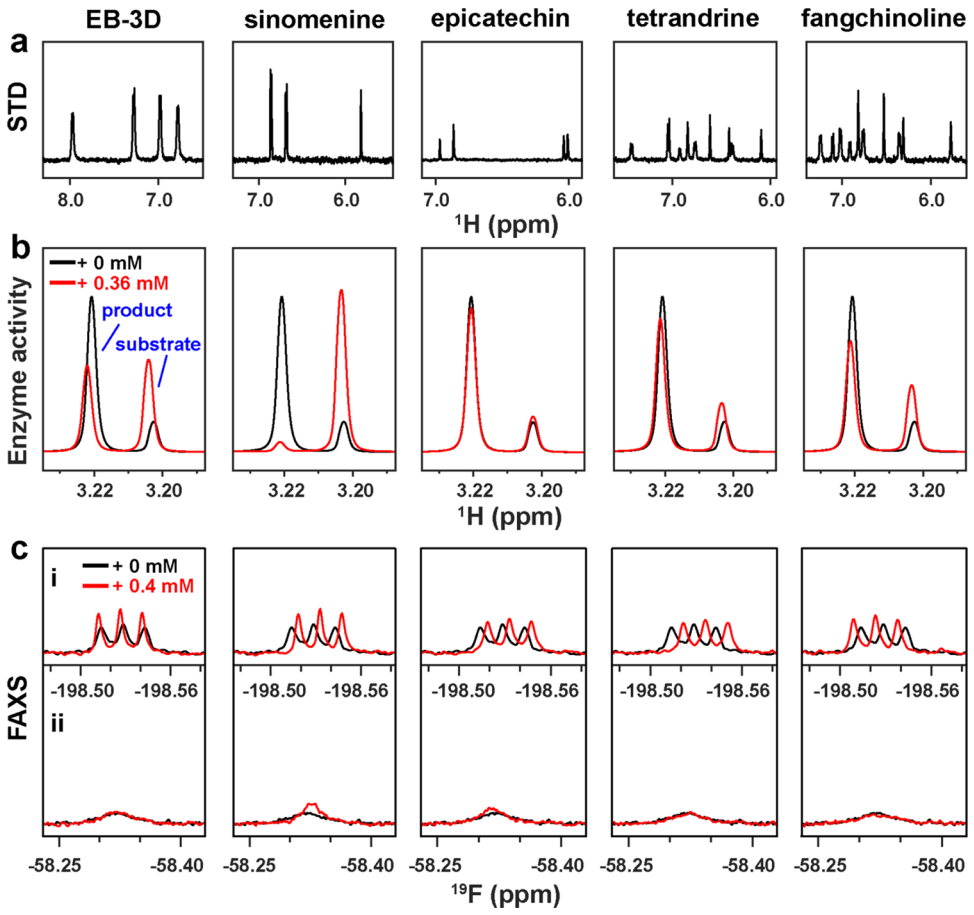


Figure 5. Series of 1D spectra of ChoKα1 inhibitors. Panel a presents the STD NMR spectra acquired with 128 scans. Panel b depicts the enzymatic reaction mixture spectra for constant-time analysis, where red lines represent the spectra with inhibitors present, and black lines show spectra without inhibitors. Peaks at 3.203 and 3.221 ppm are attributed to choline (substrate) and phosphocholine (product), respectively. Panel c displays the FAXS spectra; subpanel (i) features the NMR signals for fluorocholine chloride, while subpanel (ii) shows those for 2-fluoroadenosine.

to only 39 compounds. After comparing HSQC peak tables of these candidate compounds to the edited psHSQC peaklist of *S. Caulis* extract, sinomenine emerged with a score of 0.75, which was significantly higher than the scores of other compounds, confirming it as the active component (see Table S1 for scoring details).

Using similar approach, we successfully identified the structural motif CH_2CHCH in the active component of *C. orbiculatus* extract (Figure S5) and (−)-epicatechin as the active component (Table S2). Additionally, from the total alkaloid extract of *S. tetrandra*, two active components, tetrandrine and fangchinoline, have been identified and confirmed through spiking experiments (as shown in Figure S6 and Tables S3, S4).

Chemical shifts, subjected to solvent type, salt content, pH, and temperature variance, require careful referencing for accuracy. For example, using TMS or residual protonated solvent as reference can introduce notable disparities. To avoid false-negative results, our filtering process incorporated a wide threshold range for characteristic C−H peaks; the filtering ranges for ^{1}H and ^{13}C are 0.35 and 3.5 ppm, respectively. However, in scoring, a tight tolerance range of 0.175 ppm for ^{1}H and 1.75 ppm for ^{13}C was used to assess the likelihood of a candidate being the active component of the herbal extract. Notably, we set the ^{13}C chemical shift tolerance that is 10 times that of ^{1}H due to the less pronounced influence of environmental factors on the chemical shit of ^{13}C nuclei.[51]

The above case demonstrates the ability of our proposed identification methodology to overcome the influence of NMR solvent and other external factors, an important factor since the bulk of NMR data is often not available in DMSO-d_6 solvent.

Specifically, after identifying the active components of herbal medicines through targeted screening, quantification can be achieved using techniques such as pure shift and $HSQC_0$ methods.[52,53] These methods provide accurate quantification of active ingredients, offering new approaches for the quality control of herbal medicines.

Verification of Active Components. After identifying the chemical structures of active components (Figure S7), we procured the compounds commercially. We then validated their interaction with ChoKα1 via STD experiments (Figure 5a) and estimated their dissociation constant K_D values (Table 1). The inhibitory activity against ChoKα1 was also assessed using constant-time reaction analysis, with Figure 5b illustrating the inhibitory effects of four novel ChoKα1 inhibitors alongside the reference inhibitor EB-3D. As depicted in it, each addition of sinomenine, epicatechin, tetrandrine, fangchinoline, and EB-3D caused attenuation of the peak at 3.221 ppm, relative to the blank control, indicating a reduction in phosphocholine production and affirming their role as ChoKα1 inhibitors. Further investigation into the interactions between the active components and ChoKα1 was carried out

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

Table 1. New ChoKα1 Inhibitors Discovered from Herbal Extracts

no.	name	K_D[a]	binding site[b]
1	sinomenine	30 ± 9 μM	choline and ATP
2	epicatechin	>1 mM	choline
3	tetrandrine	229 ± 98 μM	choline
4	fangchinoline	498 ± 72 μM	choline
5	EB-3D[c]	42 ± 9 μM	choline

[a]K_D were measured via STD NMR experiments under conditions with 6% DMSO (v/v), 50 mM sodium phosphate, and 100 mM NaCl in the buffer. For a comparison with ITC measurements, see Table S5 in the Supporting Information. [b]Binding sites were determined by NMR-based competition experiments. [c]EB-3D is a known inhibitor used as a positive control in the assays.

through competitive binding experiments using spy molecules, namely 2-fluoroadenosine and fluorocholine chloride, to probe the ATP and substrate binding sites, respectively, utilizing FAXS. Figure 5c depicts the competitive experiments of four new ChoKα1 inhibitors and the reference inhibitor EB-3D. The introduction of 20 μM ChoKα1 precipitated a marked decrease in signal intensity for both spy molecules at −52.5 ppm for 2-fluoroadenosine and −192.7 ppm for fluorocholine chloride, indicating successful binding to ATP and choline sites of ChoKα1, respectively. The inhibitors at a concentration of 400 μM could significantly restore the NMR signals by displacing the spy molecules from the respective binding sites. Table 1 summarizes the four active components along with their corresponding binding sites.

Structural analysis identified three distinct chemical backbones among the components: morphine alkaloid, flavonoids, and bis-benzylisoquinoline alkaloid as shown in Figure S7. These structurally novel skeletons demonstrated effective disruption of ChoKα1 enzymatic activity, suggesting their potential for developing robust inhibitors. All four components showed competitive binding to the choline site, while, sinomenine also engaged competitively with the ATP site, validating the screening methodology's capability to identify active moieties that target disparate binding sites on ChoKα1.

Tetrandrine and fangchinoline, as bis-benzylisoquinoline alkaloids, displayed strikingly similar structures, differentiated by only a methyl group. Both selectively engaged the choline-binding site on ChoKα1, with closely matched K_D values, suggesting chemical structure congruence correlates with consistent inhibitory activity and NMR spectral shift proximity. Such structural resemblance explains the identical ranking ascribed to tetrandrine and fangchinoline during the structural identification phase within the alkaloid extract of *S. tetrandra*. Within a given herbal matrix, its chemical constituents usually belong to only a few prevalent types of skeletons. Although these components may vary significantly in concentration, they often share related structural features, leading to analogous high scores in the structural identification phase. However, a high score does not necessarily denote active components but rather a high likelihood of sharing a structural skeleton with an active moiety. These findings endorse the NMR-based drug screening strategy as a potent tool for exploring herbal extracts in search of pharmacologically active entities, even in modest concentrations.

CONCLUSIONS

The targeted screening and identification of active ingredients in herbal medicine play a crucial role in lead discovery for drug development. Here, we have introduced a streamlined method allowing efficient screening and identification of active ingredients in herbal extracts. Our approach uses peCPMG spectra to investigate herbal extract samples in the presence and absence of the target protein. This process effectively isolates the characteristic ^{1}H signals of active components. Once these signals are pinpointed, we employed the edited psHSQC and H2BC spectra to discern the characteristic C−H signals and CH_n structural motifs of the active components. These specific signals serve as screening filters for potential candidates within a specially designed database.

We enhanced the accuracy of candidate identification by aligning the C−H chemical shits of each database's candidate with the peaklist from the edited psHSQC spectrum of the herbal extract. The matching scores of candidates indicate the likelihood of their structures corresponding to the active components under investigation. This matching score prioritizes candidates, leading to the deduction of the true chemical structure of the active components, thus bypassing the laborious and time-consuming steps of isolation and structural determination traditionally required in the discovery of active components.

In summary, our study presents an integrated screening protocol that combines ligand-detected NMR with a customized database for the swift identification of active components in herbal extracts without necessitating the purification of discrete chemical entities. Validated using ChoKα1 and its known inhibitors, our technique has proven to effectively detect ligands with both strong and weak affinities across a diverse concentration spectrum. The practical application of the proposed method has not only identified novel ChoKα1 inhibitors but has also led to their confirmation via STD NMR assays. Identifying four previously unknown inhibitors associated with three novel structural backbones highlights the efficacy of our approach. In light of these findings, our method shows substantial promise for widespread application in precisely screening active components within complex herbal matrices.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.4c02255.

An example of the data structure, additional experimental procedures, tables, and figures (PDF)

AUTHOR INFORMATION

Corresponding Authors

Maili Liu − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China; Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430071, China; Optics Valley Laboratory, Wuhan 430074, China;*

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

orcid.org/0000-0002-9359-915X; Email: ml.liu@wipm.ac.cn

Xu Zhang − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China; Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430071, China; Optics Valley Laboratory, Wuhan 430074, China;* orcid.org/0000-0003-2481-946X; Email: zhangxu@wipm.ac.cn

Authors

Tao Huang − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China*

Xin Chai − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China*

Shuangli Li − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China*

Biao Liu − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430071, China*

Jianhua Zhan − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China*

Xiaohua Wang − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China*

Xiong Xiao − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China*

Qinjun Zhu − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China*

Caixiang Liu − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China*

Danyun Zeng − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China*

Bin Jiang − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China; Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430071, China; Optics Valley Laboratory, Wuhan 430074, China*

Xin Zhou − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China; Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430071, China; Optics Valley Laboratory, Wuhan 430074, China;* orcid.org/0000-0002-5580-7907

Lichun He − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China*

Zhou Gong − *Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National*

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204

Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China; University of Chinese Academy of Sciences, Beijing 100049, China

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.4c02255

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by National Key R&D Program of China (2023YFA1607500, 2018YFA0704002, 2018YFE0202300), National Natural Science Foundation of China (22327901, 22174152, 21505153, 21475146, 21991081, 21735007, and 32071990), Hubei Provincial Natural Science Foundation of China (2023AFA041), and Strategic Priority Research Program of the Chinese Academy of Sciences (XDB0540000).

■ REFERENCES

(1) Lin, L.; Luo, L.; Zhong, M.; Xie, T.; Liu, Y.; Li, H.; Ni, J. *RSC Adv.* **2019**, *9*, 17457−17472.
(2) Khazir, J.; Mir, B. A.; Mird, S. A.; Cowan, D. *J. Asian Nat. Prod. Res.* **2013**, *15*, 764−788.
(3) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2020**, *83*, 770−803.
(4) Moaddel, R.; Marszałł, M.; Bighi, F.; Yang, Q.; Duan, X.; Wainer, I. *Anal. Chem.* **2007**, *79*, 5414−5417.
(5) Huang, Q.; Gao, Q.; Chai, X.; Ren, W.; Zhang, G.; Kong, Y.; Zhang, Y.; Gao, J.; Lei, X.; Ma, L. *J. Chromatogr. B* **2020**, *1151*, No. 122153.
(6) Tao, Y.; Zhang, Y.; Cheng, Y.; Wang, Y. *Biomed. Chromatogr.* **2013**, *27*, 148−155.
(7) Zhang, B.; Zhao, S.; Yang, D.; Wu, Y.; Xin, Y.; Cao, H.; Huang, X.-P.; Cai, X.; Sun, W.; Ye, N.; Xu, Y.; Peng, Y.; Zhao, S.; Liu, Z.-J.; Zhong, G.; Wang, M.-W.; Shui, W. *ACS Cent. Sci.* **2020**, *6*, 213−225.
(8) van Breemen, R. B.; Huang, C.-R.; Nikolic, D.; Woodbury, C. P.; Zhao, Y.-Z.; Venton, D. L. *Anal. Chem.* **1997**, *69*, 2159−2164.
(9) Comess, K. M.; Schurdak, M. E.; Voorbach, M. J.; Coen, M.; Trumbull, J. D.; Yang, H.; Gao, L.; Tang, H.; Cheng, X.; Lerner, C. G.; et al. *J. Biomol. Screening* **2006**, *11*, 743−754.
(10) Song, H.-P.; Chen, J.; Hong, J.-Y.; Hao, H.; Qi, L.-W.; Lu, J.; Fu, Y.; Wu, B.; Yang, H.; Li, P. *Chem. Commun.* **2015**, *51*, 1494−1497.
(11) An, Y.; Li, X.; Sun, H.; Bian, W.; Li, Z.; Zhang, Y.; Zhao, X.; Zheng, X. *J. Mol. Recognit.* **2015**, *28*, 628−634.
(12) Chen, X.; Cao, Y.; Zhang, H.; Zhu, Z.; Liu, M.; Liu, H.; Ding, X.; Hong, Z.; Li, W.; Lv, D.; et al. *Anal. Chem.* **2014**, *86*, 4748−4757.
(13) Kaur, S.; McGuire, L.; Tang, D.; Dollinger, G.; Huebner, V. *J. Protein Chem.* **1997**, *16*, 505−511.
(14) Wang, L.; Gao, J.; Ma, R.; Liu, Y.; Liu, M.; Zhong, F.; Hu, J.; Li, S.; Wu, J.; Jiang, H.; et al. *Magn. Reson. Lett.* **2022**, *2*, 107−118.
(15) Sugiki, T.; Furuita, K.; Fujiwara, T.; Kojima, C. *Molecules* **2018**, *23*, 148.
(16) Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784−1788.
(17) Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M. *J. Biomol. NMR* **2000**, *18*, 65−68.
(18) Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. *J. Biomol. NMR* **2001**, *21*, 349−359.
(19) Sun, P.; Jiang, X.; Jiang, B.; Zhang, X.; Liu, M. *J. Biomol. NMR* **2013**, *56*, 285−290.
(20) Bingol, K.; Li, D.-W.; Bruschweiler-Li, L.; Cabrera, O. A.; Megraw, T.; Zhang, F.; Brüschweiler, R. *ACS Chem. Biol.* **2015**, *10*, 452−459.
(21) Huang, T.; Chen, P.; Liu, B.; Li, X.; Lv, X.; Hu, K. *Anal. Chem.* **2020**, *92*, 10996−11006.
(22) Wan, H.; Tian, Y.; Jiang, H.; Zhang, X.; Ju, X. *J. Ethnopharmacol.* **2020**, *254*, No. 112712.
(23) Zhou, Q.; Li, L.; Xiang, J.; Tang, Y.; Zhang, H.; Yang, S.; Li, Q.; Yang, Q.; Xu, G. *Angew. Chem., Int. Ed.* **2008**, *47*, 5590−5592.
(24) Tanoli, S. A. K.; Tanoli, N. U.; Bondancia, T. M.; Usmani, S.; Kerssebaum, R.; Ferreira, A. G.; Fernandes, J. B.; Ul-Haq, Z. *Analyst* **2013**, *138*, 5137−5145.
(25) Tanoli, S. A. K.; Tanoli, N. U.; Bondancia, T. M.; Usmani, S.; Ul-Haq, Z.; Fernandes, J. B.; Thomasi, S. S.; Ferreira, A. G. *RSC Adv.* **2015**, *5*, 23431−23442.
(26) Wishart, D. S.; Sayeeda, Z.; Budinski, Z.; Guo, A.; Lee, B. L.; Berjanskii, M.; Rout, M.; Peters, H.; Dizon, R.; Mah, R. NP.; et al. *Nucleic Acids Res.* **2022**, *50*, D665−D677.
(27) López-Pérez, J. L.; Theron, R.; del Olmo, E.; Díaz, D. *Bioinformatics* **2007**, *23*, 3256−3257.
(28) Kuhn, S.; Schlörer, N. E. *Magn. Reson. Chem.* **2015**, *53*, 582−589.
(29) Hayamizu, K. *Kagaku to Seibutsu* **2011**, *49*, 250−255.
(30) Xia, J.; Bjorndahl, T. C.; Tang, P.; Wishart, D. S. *BMC Bioinf.* **2008**, *9*, 507.
(31) Dubey, A.; Rangarajan, A.; Pal, D.; Atreya, H. S. *Anal. Chem.* **2015**, *87*, 7148−7155.
(32) Sorokina, M.; Merseburger, P.; Rajan, K.; Yirik, M. A.; Steinbeck, C. *J. Cheminf.* **2021**, *13*, No. 2.
(33) Weininger, D.; Weininger, A.; Weininger, J. L. *J. Chem. Inf. Model.* **1989**, *29*, 97−101.
(34) Daylight Chemical Information Systems, 3. SMILES - A simplified chemical language. 2022, https://www.daylight.com/dayhtml/doc/theory/theory.smiles.html (accessed March 14, 2024).
(35) Daylight Chemical Information Systems, 4. SMARTS - A Language for describing molecular patterns. 2022, https://www.daylight.com/dayhtml/doc/theory/theory.smarts.html (accessed March 14, 2024).
(36) Lucas, L. H.; Larive, C. K.; Wilkinson, P. S.; Huhn, S. *J. Pharm. Biomed. Anal.* **2005**, *39*, 156−163.
(37) Adams, R. W.; Holroyd, C. M.; Aguilar, J. A.; Nilsson, M.; Morris, G. A. *Chem. Commun.* **2013**, *49*, 358−360.
(38) Wang, J.; Zhang, X.; Sun, P.; Jiang, X.; Jiang, B.; Cao, C.; Liu, M. *J. Magn. Reson.* **2010**, *206*, 205−209.
(39) Meyer, N. H.; Zangger, K. *Angew. Chem., Int. Ed.* **2013**, *52*, 7143−7146.
(40) Nyberg, N. T.; Duus, JØ.; Sørensen, O. W. *J. Am. Chem. Soc.* **2005**, *127*, 6154−6155.
(41) Angulo, J.; Enríquez-Navas, P. M.; Nieto, P. M. *Chem. - Eur. J.* **2010**, *16*, 7803−7812.
(42) Mariotto, E.; Bortolozzi, R.; Volpin, I.; Carta, D.; Serafin, V.; Accordi, B.; Basso, G.; Navarro, P. L.; López-Cara, L. C.; Viola, G. *Biochem. Pharmacol.* **2018**, *155*, 213−223.
(43) Lacal, J. C.; Campos, J. M. *Mol. Cancer Ther.* **2015**, *14*, 31−39.
(44) Gaudry, A.; Pagni, M.; Mehl, F.; Moretti, S.; Quiros-Guerrero, L.-M.; Cappelletti, L.; Rutz, A.; Kaiser, M.; Marcourt, L.; Queiroz, E. F.; Ioset, J.-R.; Grondin, A.; David, B.; Wolfender, J.-L.; Allard, P.-M. *ACS Cent. Sci.* **2024**, *10*, 494−510.
(45) Mazarico, J. M.; Lobo, V. J. S.-A.; Favicchio, R.; Greenhalf, W.; Costello, E.; Carrillo-de Santa Pau, E.; Marqués, M.; Lacal, J. C.; Aboagye, E.; Real, F. X. *Mol. Cancer Ther.* **2016**, *15*, 323−333.
(46) Urick, A. K.; Calle, L. P.; Espinosa, J. F.; Hu, H.; Pomerantz, W. C. *ACS Chem. Biol.* **2016**, *11*, 3154−3164.
(47) Zech, S. G.; Kohlmann, A.; Zhou, T.; Li, F.; Squillace, R. M.; Parillon, L. E.; Greenfield, M. T.; Miller, D. P.; Qi, J.; Thomas, R. M.; Wang, Y.; Xu, Y.; Miret, J. J.; Shakespeare, W. C.; Zhu, X.; Dalgarno, D. C. *J. Med. Chem.* **2016**, *59*, 671−686.

(48) Cala, O.; Krimm, I. *J. Med. Chem.* **2015**, *58*, 8739−8742.
(49) Reid, D. P. *A Handbook of Chinese Healing Herbs: An Easy-to-Use Guide to 108 Chinese Medicinal Herbs and Dozens of Prepared Herba Formulas*; Shambhala Publications: Boulder, 1995.
(50) Di, L.; Kerns, E. H. *Drug Discovery Today* **2006**, *11*, 446−451.
(51) Felli, I. C.; Pierattelli, R. *Chem. Rev.* **2022**, *122*, 9468−9496.
(52) Foster, H. M.; Nilsson, M.; Adams, R. W.; Morris, G. A. *Anal. Chem.* **2024**, *96*, 9601−9609.
(53) Hu, K.; Westler, W. M.; Markley, J. L. *J. Am. Chem. Soc.* **2011**, *133*, 1662−1665.

https://doi.org/10.1021/acs.analchem.4c02255
Anal. Chem. 2024, 96, 15194−15204