jove

NMR-based Fragment Screening in a Minimum Sample but Maximum Automation Mode

Hannes Berg[1], **M. A. Wirtz Martin**[1], **A. Niesteruk**[1,2], **C. Richter**[1], **S. Sreeramulu**[1], **H. Schwalbe**[1,2]

[1] Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance (BMRZ), Johann Wolfgang Goethe-University Frankfurt [2] German Cancer Consortium (DKTK) and DKFZ

[*] These authors contributed equally

Corresponding Author

H. Schwalbe
schwalbe@nmr.uni-frankfurt.de

Abstract

Fragment-based screening (FBS) is a well-validated and accepted concept within the drug discovery process both in academia and industry. The greatest advantage of NMR-based fragment screening is its ability not only to detect binders over 7-8 orders of magnitude of affinity but also to monitor purity and chemical quality of the fragments and thus to produce high quality hits and minimal false positives or false negatives. A prerequisite within the FBS is to perform initial and periodic quality control of the fragment library, determining solubility and chemical integrity of the fragments in relevant buffers, and establishing multiple libraries to cover diverse scaffolds to accommodate various macromolecule target classes (proteins/RNA/DNA). Further, an extensive NMR-based screening protocol optimization with respect to sample quantities, speed of acquisition and analysis at the level of biological construct/fragment-space, in condition-space (buffer, additives, ions, pH, and temperature) and in ligand-space (ligand analogues, ligand concentration) is required. At least in academia, these screening efforts have so far been undertaken manually in a very limited fashion, leading to limited availability of screening infrastructure not only in the drug development process but also in the context of chemical probe development. In order to meet the requirements economically, advanced workflows are presented. They take advantage of the latest state-of-the-art advanced hardware, with which the liquid sample collection can be filled in a temperature-controlled fashion into the NMR-tubes in an automated manner. ^{1}H/^{19}F NMR ligand-based spectra are then collected at a given temperature. High-throughput sample changer (HT sample changer) can handle more than 500 samples in temperature-controlled blocks. This together with advanced software tools speeds up data acquisition and analysis. Further, application of screening routines on protein and RNA samples are described to make aware of the established protocols for a broad user base in biomacromolecular research.

Introduction

Fragment-based screening is now a commonly used method for identifying rather simple and low molecular weight molecules (MW <250 Da) that show weak binding to macromolecular targets including proteins, DNA and RNA. Initial hits from primary screens serve as the basis to conduct a secondary screen of commercially available larger analogues of the hits and then to utilize chemistry-based fragment growth or linking strategies. For a successful fragment-based drug discovery (FBDD) platform, in general, a robust biophysical method is required for detecting and characterizing weak hits, a fragment library, a biomolecular target and a strategy for follow-up chemistry. Four commonly applied biophysical methods within the drug discovery campaigns are thermal shift assays, surface plasmon resonance (SPR), crystallography and nuclear magnetic resonance spectroscopy (NMR).

NMR spectroscopy has displayed varied roles within the different stages of the FBDD. Apart from ensuring the chemical purity and solubility of the fragments in a fragment library dissolved in an optimized buffer system, ligand observed NMR experiments can detect fragment binding to a target with low affinity and the target observed NMR experiments can delineate the binding epitope of the fragment, thus enabling detailed structure-activity relationship studies. Within epitope mapping, NMR-based chemical shift changes cannot only identify the orthosteric binding sites but also allosteric sites that might be cryptic and only accessible in so called excited conformational states of the biomolecular target. If the biomolecular target already binds an endogenous ligand, the identified fragment hits can be easily classified as allosteric or orthosteric by performing NMR-based competition experiments. Determining the dissociation constant (K_D) of the ligand-target interaction

is an important aspect in the FBDD process. NMR-based chemical shift titrations, either ligand or target observed can be readily performed to determine the K_D. A major advantage of NMR is that the interaction studies are performed in solution and near to physiological conditions. Thus, all conformational states for the analysis of ligand/fragment interaction with its target can be probed. Further, NMR-based approaches are not only restricted towards screening of well-folded soluble proteins, but also are being applied to accommodate larger target space including DNA, RNA, membrane bound and intrinsically disordered proteins[1].

Fragment libraries are an indispensable part of the FBDD process. In general, fragments act as the initial precursors which eventually become part (substructure) of the new inhibitor developed for a biological target. Several drugs (Venetoclax[2], Vemurafenib[3], Erdafitinib[4], Pexidartnib[5]) have been reported to have started as fragments and are now successfully used in the clinics. Typically, fragments are low molecular weight (<250 Da) organic molecules with a high aqueous solubility and stability. A carefully crafted fragment library containing typically a few hundred fragments, can already promise efficient exploration of chemical space. The general composition of fragment libraries has evolved overtime and most often were derived by dissecting known drugs into smaller fragments or designed computationally. These diverse fragment libraries mainly contain flat aromatic or heteroatoms and adhere to the *Lipinski Rule of 5* [6]*,* or to the current commercial trend *Rule of 3* [7]*,* but avoid reactive groups. Some fragment libraries were also derived or composed of highly soluble metabolites, natural products and or their derivatives[8]. A general challenge posed by most of the fragment libraries is ease of downstream chemistry.

The Center for Biomolecular Magnetic Resonance (BMRZ) at the Goethe-University Frankfurt, is a partner of the iNEXT-Discovery (Infrastructure for NMR, EM and X-rays for Translational research-Discovery), a consortium for structural research infrastructures for all European researchers from all fields of biochemical and biomedical research. Within the previous initiative of iNEXT which ended in 2019, a fragment library comprising of 768 fragments was crafted with the aim of "minimum fragments and maximum diversity" covering a large chemical space. Further, unlike any other fragment library, the iNEXT fragment library was also designed based on the concept of "poised fragments" with the aim to ease downstream synthesis of complex, high affinity ligands and henceforth known as in-house library (Diamond, Structural Genomic Consortium and iNEXT).

Establishing FBDD by NMR requires manpower, knowledge, and instrumentation. At the BMRZ, optimized workflows to support technical assistance to fragment screening by NMR have been developed. These include quality control and solubility assessment of the fragment library [9], buffer optimization for the chosen targets, ^{1}H or [19]F- observed 1D-ligand based screening, competition experiments to differentiate between orthosteric and allosteric binding, 2D-based target observed NMR experiments for epitope mapping, and for characterizing the interaction with secondary set of derivatives of the initial fragment hits. BMRZ has established automated routines for the analysis, as also previously discussed in the literature [10,11], of small molecule-protein interactions and has in place all the necessary automated infrastructure for NMR-based fragment screening. It has implemented saturation transfer difference NMR (STD-NMR), water-ligand observed via gradient spectroscopy (waterLOGSY), and Carr-Purcell-Meiboom-Gill-based (CPMG-based) relaxation experiments

to identify fragments within a wide range of affinity regimes as well as state-of-the-art automated NMR instrumentation and software for drug discovery. While NMR-based fragment screening is well established for proteins, this approach is less commonly used for finding new ligands interacting with RNA and DNA. BMRZ has established proof of concept for new protocols enabling the identification of small molecule-RNA/DNA interactions. In the following sections of this contribution, application of screening routines on protein and RNA samples is reported to make aware of the established protocols for a broad user base in biomacromolecular research.

Protocol

1. Fragment library

1. In-house fragment library
   NOTE: Within the framework of one of the joint research activities of the iNEXT, a robust and downstream chemistry friendly first generation fragment library was developed[12] and subsequently a second generation of the library was put together in collaboration with Enamine and is known as the DSI (Diamond-SGC-iNEXT)-poised fragment library (from now on termed as "In-house library"). This library can be made available at the BMRZ for screening purposes.

   1. Assess the fragment library for its integrity and solubility using a previously reported NMR-based protocol[9].
      NOTE: The in-house library consists of 768 fragments with a very high chemical diversity (>200 Singletons). Performing the screening in fragment mixtures can significantly speed up the screening campaign; however, the number of fragments in a mix is limited due to signal overlap in ^{1}H-NMR

spectrum. The higher chemical diversity offered by the in-house library allows for the preparation of mixtures containing 12 different fragments without any significant chemical shift overlap in the ^{1}H observed NMR spectra.

2. 103 fragments within the 768 fragments possess a fluorine atom. For ^{19}F screening purposes, divide all 103 fragments that possess a fluorine group into 5 mixes based on minimum ^{19}F chemical shift overlap. To minimize signal overlap in the ^{19}F screening, use the chemical shift information from single compound measurements to design mixtures with maximum number of fragments and minimal signal overlap. Each mix has 20-21 fragments with distinct ^{19}F chemical shifts allowing for unambiguous assignment of fragments.

2. User defined/provided fragment library

1. Perform screening campaigns with the user defined or provided fragment library; however, the following steps need to precede the screening campaign.

2. If not specified by the user beforehand, perform NMR-based quality control of the fragments (at the BMRZ, advanced software tools are used for this; [9], Chapter 6.1.1).

3. Check the solubility of the fragments in buffer-of-choice for the biomolecular target, structure integrity, and concentration of fragments prior to use.

4. Design the mixture to decrease both signal overlap in NMR spectra and measurement time.

5. Design mixtures according to step 4.2.

6. Screen single fragments or a subset of mixtures instead of the entire library.

2. Sample preparation

NOTE: High-throughput screening by NMR utilizes a pipetting robot for sample preparation. NMR spectra, but also stabilities over several days of signal acquisition of proteins, RNAs and DNA are extremely sensitive to temperature fluctuations and therefore temperature-controlled automated systems will greatly facilitate the stability of the samples being pipetted. For this purpose, an additional add-on device, which works between 4 to 40 °C, is coupled to the pipetting robot for liquid handling of the NMR samples in a temperature-controlled environment.

1. Ligand mixture preparation

1. Prepare screening samples for NMR measurements using a sample preparation robot. The flexible configuration of the robot allows for a wide range of applications (e.g., recovery of the samples from NMR tubes back into storage containers or general liquid handling tasks). NMR-tubes with different diameters (1.7, 2.0, 2.5, 3.0 and 5.0 mm) can be used. The sample robot system along with the advanced control software reads the barcode assigned for each container type and executes the liquid filling protocol optimally.

2. For the preparation of the in-house library ligand mixtures, use barcoded vials. The barcoded vials guarantee the highest level of reliability and optimal traceability of the samples.

3. Distribute 768 compounds into 8 plates of 96-well format. The stock concentration of each individual fragment is 50 mM in d$_6$-DMSO/D$_2$O (9:1). In total, prepare 64 mixes each containing 12 fragments. The

final concentration of each fragment in a mix is 4.2 mM.

NOTE: The pipetting robot can accommodate a variety of container types with varied geometries (cryo- or auto sampler vials, 96-well plates round or square deep, barcoded standard vials, microcentrifuge tubes) and assists efficient execution of the liquid transfer to a variety of NMR tubes and racks.
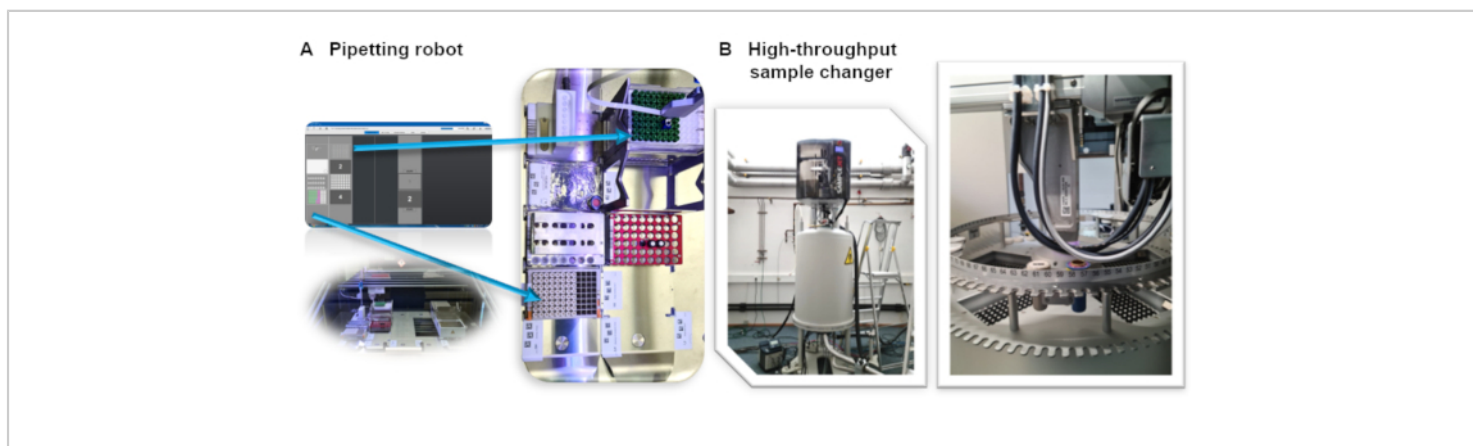
Figure 1: (**A**) High-throughput NMR sample preparation and NMR-tube filling robot installed at BMRZ. (**B**) High-throughput sample changer with individual temperature-controlled racks installed on a 600 MHz spectrometer at the BMRZ facility. Please click here to view a larger version of this figure.

2. Screening sample preparation blank (reference ligand spectrum) and with target (ligand in the presence of target)

   1. For the preparation of the NMR screening samples, in the presence of the target biomolecule (protein/RNA/DNA) and the ligand mixture, use 3 mm NMR HT sample changer tubes selected from the Bruker NMR portfolio of standard NMR tubes.

   2. Transfer the biomolecular target (e.g., ^{1}H Screening: 10 µM RNA or Protein) in a defined screening buffer into the 3 mm NMR tube (final volume of 200 µL) manually or using the pipetting robot.

3. Transfer 10 µL (e.g., ^{1}H Screening) of the ligand mixture in the next step using the robotic system into the barcoded 3 mm NMR tubes containing the target biomolecule and mix using the inbuilt protocol of the control software.

   NOTE: The barcode number of the NMR tube is conveniently and automatically incorporated into the acquired NMR-dataset, thus ensuring ID oriented workflow without any mix-up. The pipetting robot temperature control accessory allows to keep the prepared samples in the NMR tubes under constant temperature.

3. In-house defined conditions and parameters

   1. Establish optimal buffer conditions for performing screening of RNA and protein against the in-house

fragment library. The following sample conditiond are used for the RNA at the BMRZ: 25 mM KPi, 50 mM KCl, pH 6.2. Mg^{2+} is optional.

2. Proteins are extremely sensitive for solution conditions; use buffers optimal for the target of choice. For each of those buffers, acquire additional reference spectra of the ligands to serve as blank for the analysis.

4. User specified conditions

NOTE: In cases in which the in-house established conditions are not suitable for the targets to be screened from a potential user, the following steps should be implemented.

1. Perform ^{1}H-NMR of the buffer alone to ensure minimal interference from the components of the buffer in performing and analysing the ligand observed screening experiments. Interfering components could be suitably replaced with deuterated equivalents.

5. Limitations in sample production (target quantities)/ conditions and availability

NOTE: Isolation or recombinant production of certain biomacromolecules can in certain cases prove challenging and result in limited availability of the target to pursue a successful drug screening campaign. In cases of limited or unlimited availability of the targets, the following alternatives could be utilized for conducting a successful NMR-based fragment screening.

1. If limited, use ^{19}F-NMR based screening. Typical fluorinated ligands have a single ^{19}F signal; therefore, use cocktails with 25-30 fragments without any signal overlap. There are fewer signals to analyse, no signal interference from buffer

components, and fewer signals to rely on for hit identification.

2. If unlimited, use larger screens like ^{1}H-NMR. The larger fragment library can be screened. Typically, fragments are composed of more than one proton, which means more signals to rely upon for the analysis.

3. NMR acquisition conditions

1. In-house generally defined conditions

   1. Spectrometer equipped with HT sample changer (Automation)

      1. For high-throughput screening, use 96-well plates that can only be measured using a HT sample changer. The HT sample changer also offers the possibility to temper each rack individually.

      2. For optimal signal-to-noise, use a spectrometer with a cryogenic probe that is either helium or nitrogen cooled. An automated tunning and matching module (ATM) are necessary for automation.

   2. Parameter sets & pulse sequences

      NOTE: Many NMR experiments can characterize binding events. The hit identification varies depending on the experimental setup. The following experiments are routinely used in BMRZ screening campaigns. Nevertheless, changes can be made for user defined screening campaigns and according to user specifications.

      1. If usingTopSpin software, include the parameter set for ligand-based experiments: SCREEN_STD, SCREEN_T1R, SCREEN_T2,

SCREEN_WLOGSY. The parameter set includes all necessary parameters and the pulse sequences: STD: stddiffesgp.3; $T_{1\rho}$: t1rho_esgp2d; T_2: cpmg_esgp2d; and waterLOGSY: ephogsygpno.2.

2. For all the listed experiments, use excitation sculpting[13] as water suppression. For a reference, use 1D excitation sculpting (zgesgp). The number of scans depends on the sensitivity of the system (magnetic field strength and probe head), the sample concentration, and choice of the experiment. A recommendation is: 1D with NS=64, $T_{1\rho}$ & T_2 with NS=128, STD with NS=256 and waterLOGSY with NS= 384 or 512.

3. For the ^{19}F screening, use both 1D and T_2 experiments: 1D: F19CPD (pp=zgig) for ^{19}F$\{^1$H$\}$-probe head and F19(pp=zg) for ^{19}F/^{1}H-probe head; SCREEN_19F_T2 (pp = cpmgigsp).

4. Use a spectral width of 220 ppm and an excitation frequency at -140 ppm. The experiment time is between 1 and 5 hours (ensure the long-term stability of the biomacromolecule) depending on the hardware and sample concentration. For T_2, the CPMG time should alternate between 0 ms and 200 ms.

3. Processing

1. Record the STD, $T_{1\rho}$ and T_2 experiments as pseudo 2D. To process the two single 1D spectra, IconNMR uses the au-program proc_std either with or without the option relax.

The first option provides the reference 1D and the difference of two spectra. The second option yields two separate spectra with short and long relaxation time. The waterLOGSY is a single 1D which should be phased with a negative for the solvent signal.

2. User specific conditions

1. Adapt any of the previously mentioned parameters to user-defined conditions. For example, if a facility user-provided protein is not stable at the generally used temperature, optimization experiments can be conducted varying temperature, concentration, buffer conditions etc.

4. Data Analysis

1. Fragment library QC (d6-DMSO/specific buffer) and quantification

1. CMC-q

NOTE: Quality control of fragment libraries is essential prior to initiation of screening campaigns. Furthermore, long-term stability of fragment library needs to be ensured for the application of several screening campaigns, which is why periodic evaluation of the quality of the library must be conducted. For this purpose, the integrated software CMC-q and CMC-a from TopSpin is used for quality and quantity assessment. The CMC-q and CMC-a are software modules within Topspin which enable smooth acquisition, analysis including structure verification using ^{1}H-NMR spectrum obtained from small organic molecules [9].

1. For integrity, prepare assessment samples with a fragment concentration of 1 mM in d6-DMSO.

Prepare samples in an automated manner with a pipetting robot by filling liquid sample collection into a 3 mm NMR-tube.

2. For solubility assessment, use sample consisting of 1 mM compound in 50 mM sodium phosphate buffer at pH 7.4, 150 mM sodium chloride, 90% H_2O/ 10% D_2O and 1 mM of 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid sodium salt (TMSP-Na).

3. Collect NMR spectra at 298 K or 293 K using a 600 MHz NMR spectrometer equipped with triple resonance 5 mm TCI cryogenic probe and a HT sample changer, which can handle 579 samples at once.

4. For setting up CMC-q software follow the instructions of the user manual, which implements the creation of an IconNMR user, the activation of FastLaneNMR, and changing the HT sample changer.

5. Calibrate the 90° pulse and save it in the TopSpin prosol table.

6. Place the 96 sample well plate in one of the 5 rack positions in the HT sample changer.

7. To load an SDF file (structure data file) that should contain its proposed chemical structure, a unique identifier, and the position in the HT sample changer of each sample in a batch, go to **Browse** in the CMC-q Setup window and click **Open** after selecting a file that ends in .sdf.

8. In the CMC.q Batch Automation settings, set the verification type that defines the experiment that will be measured, the IconNMR user and define the Solvent.

9. Define SDF files for the Path for SDF file, the molecule ID and the sample position.

10. Start the acquisition by clicking on **Start**. Click on **Start Acquisition** again. The CMC-q Setup can also be saved by clicking on **Save**.

11. For in-detail description of CMC-q setup steps, follow the user manual instructions from Bruker.

2. CMC-a

1. For CMC-a, use the software module within Topspin that enables analysis including structure verification using 1H-NMR spectrum obtained from small organic molecules[9].

2. Mixture design

NOTE: A proper mixture design plays an important role for screening using NMR as a platform. A high number of fragments per mixtures allows for faster screening but increases the risk of false positive and negatives. A lower number decreases that risk but increases the time it takes to conduct the screening. In general, a signal overlap has to be avoided when creating mixtures. Using the In-house library, this can be neglected for the ^{1}H screening as the library was specifically designed to be diverse and show little signal overlap while maintaining a high chemical diversity. This in turn means that no special design procedure has to be undergone for creating the 64 mixes.

1. As the ^{19}F screening relies on the fragments of the in-house library that contain fluorine and the library was not created to reduce the signal overlap for these specific fragments, design a proper mixture.

2. Measure single compound spectra for all fragments containing [19]F.

3. Note the chemical shift information of each signal.

4. According to this information, choose 20-21 fragments per mixture. This in turn gives 5 mixtures each containing 20-21 fragments with no signal overlap and allows a semi-automated analysis of the data.

3. Perform hit identification within a ligand observed biomacromolecule-ligand interaction

NOTE: There are different definitions of a hit between the ^{19}F and ^{1}H screening procedure. The following hit identifications were set up by us and follow specific rules. The subject of hit determination is a very subjective manner and can differ from user to user. Nevertheless, it is of utmost importance that the rules for hit identification do not change once agreed upon to maintain validation and credibility.

1. **^{1}H Screen**

   1. To confidently determine hits, acquire 1D ^{1}H spectra, waterLOGSY and T_2 relaxation experiments both in the presence and absence of target to identify binders. All three experiments have the potential to show a binding event. If a CSP of greater than 6 Hz is visible in the sample spectra compared to the blank spectra, this is considered as an indication for a hit. The same goes if a strong positive signal in the waterLOGSY as well as more than 30% T_2 reduction in the sample spectra is visible. Binding events can be showcased in all three experiments, when comparing the sample containing spectra with their respective blank spectra. However binding events may not be visible in all three

experiments. Because of this it was agreed upon that at least two of the before described events must occur in order to classify a fragment as a binding hit.

2. Use the FBS tool in TopSpin to define the state of fragments into binding, ambiguous, unknown, aggregates and not-binding.

3. When finished with a mix, approve it within the FBS tool.

4. In the summary tab within the FBS project, click on **Create a screening report**. This will open a window that creates a .xlsx file. The user may then opt to choose between all ligands, binding ligand only, not binding ligand only and ambiguous ligands to be reported in the spreadsheet.

2. ^{19}F Screen

   1. To differentiate between non-binder, week-binder, and strong-binder, divide the integration quotient between the 200 ms target measurement and the 200 ms blank measurement by the quotient of the 0 ms target measurement and the 0 ms blank measurement is used:

   $$Hit - Score = \frac{\dfrac{Target^{200ms}}{Reference^{200ms}}}{\dfrac{Target^{0ms}}{Reference^{0ms}}}$$

   NOTE: This gives values ranging from 0 to ~1 (the hit-score), making it possible to assign thresholds for each binding state.

   2. Use the average of the reference 200 ms measurement as a baseline threshold, to mark

cases where the hit-score exceeds 1. This can occur, if the imported integrals contain negative values or the reference measurement is higher than the target measurement. A hit-score of ≤ 0.67 is considered a weak-hit, < 0.33 a strong hit, and anything > 0.67 as no-hit. An example is shown in **Figure 2**.
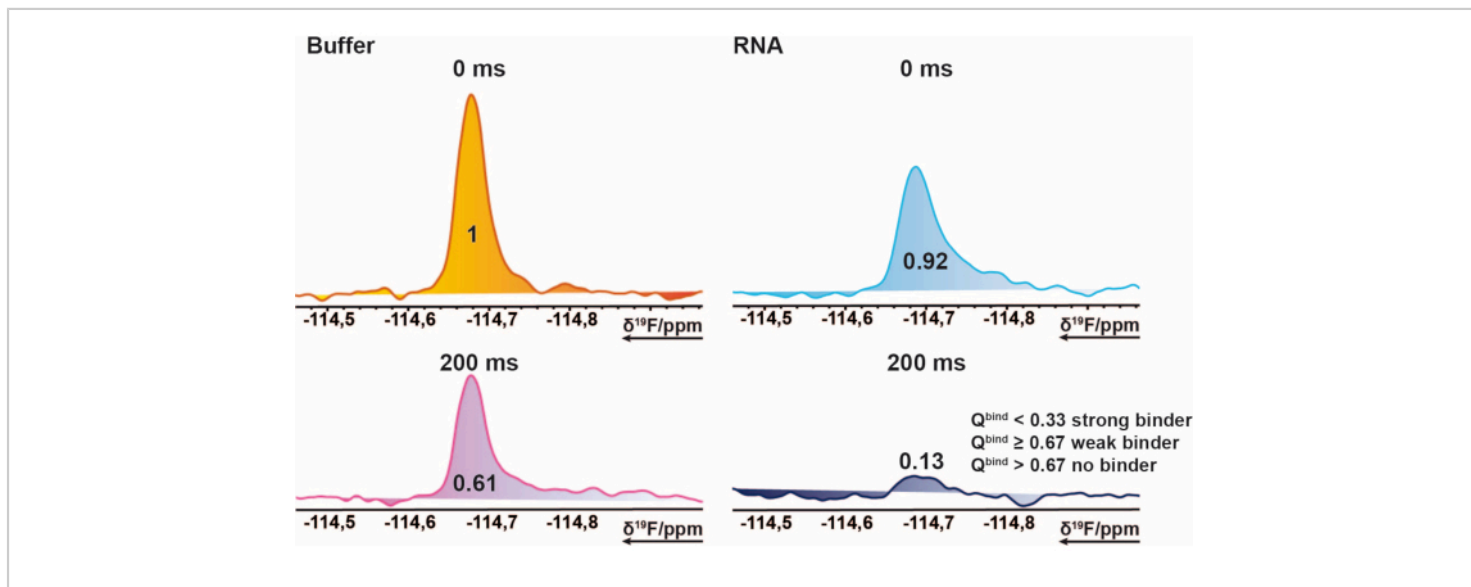
Figure 2: Hit identification for the ^{19}F screening. Section of ^{19}F CPMG NMR spectra of an exemplary compound. This pictorial representation explains the properties of a binder. ^{19}F-CPMG spectra of a compound acquired of mixture samples in the presence and absence of RNA. The values represent the normed integral values of the corresponding peak. Please click here to view a larger version of this figure.

4. Data analysis

   1. Prepare data for analysis

      NOTE: It is important that the acquired data has no visible flaws. This means that data where the shimming was problematic, or water suppression was insufficient should not be considered for analysis. Rather it is recommended to record data again and making sure that everything is fine with the sample (e.g., no air bubbles), with the temperature, the shimming, and water suppression. Data correctness can always be assessed when comparing DMSO signals.

2. ^{1}H screening

   1. To analyse ^{1}H screening data, use the FBS tool (needs additional license) in TopSpin 4.0.9.

   2. Follow the instructions in the FBS tool manual to start with the data analysis. The following steps summarize the procedure reported in the manual.

   3. Store the BMRZ NMR data from screening campaigns such that each different screening mixture has its own directory in which a

subdirectory holds the different experiments measured on the sample.

4. For using the FBS tool, store the reference spectra that has all data saved from samples without the biomolecular target but with the mixtures as well as the single compound measured in different /nmr directories. This is important as FBS tool will ask for the directory path of each individually.
NOTE: The FBS tool will recognize a directory as a screening project if the following datasets were stored in the same directory where the mixtures of a screening sample are stored (csv, FragmentScreen XML documents and BAK file).

5. When using TopSpin 4.0.9, create a direct path to the directory containing the acquired data, a so-called DIR. Choose the /nmr directory in which all mixtures should have a distinct directory.

6. To start the FBS tool of a screened sample drag the symbol FBS project into the middle of the TopSpin window. In the chosen directory the FBS project symbol should appear if previously said datasets were copied into it.

7. The window **Fragment Based Screening Options** should automatically open when first loading a new FBS project. In this window choose a cocktail file. The cocktail file is a csv file containing the assignment of the name of the mixes, the name of each fragment and their division into the mixes. Also define a reference ligand spectra folder which has all measured

spectra of the single fragments. Lastly, define a reference blank experiment folder, which is usually the folder containing the datasets of the mixes without the investigated target.

8. The Fragment Based Screening Options has a tab called **Spectra types** that lets one define the investigated spectra as well as the colour for displaying the spectra. Set the Spectype according to the beforehand processed data. In the **Display layout** tab, define the spectra that will be compared with each other according to their spectypes.

9. Press **Ok** to start the FBS project.

10. While looking at the data, a separate window will open, summarizing all cocktail mixes and all ligands of each mix in a table. By double clicking on a cell, the respective datasets will open, comparing for example ^{1}H 1D Blank spectra with the dataset containing the target.

11. Before assigning binders make sure that reference peaks (DMSO of all measurements as well as the single compounds) match with each other and have the same chemical shift. If differences are observed, correct them by using the serial processing option from TopSpin.

12. The serial processing option is under the **Process** tab under **Advanced**. It applies changes to all selected spectra from a dataset. This way, Spectypes can easily be assigned to experiment numbers and all spectra can be shifted at once to align with the reference.

3. ^{19}F Screening

1. For the first analysis of the ^{19}F mixtures, create an integration file for each mix. To define the integration region, click on the **Integrate** function in the **Analyse** tab. Make sure that for every fragment in the mixture a clear integration region for the corresponding ^{19}F-singal is defined.

2. Use the **Save/Export integration regions** button to export the integration-file for future use. Save any used integration-files in C:\Bruker\TopSpin4.0.9\exp\stan\nmr\lists\intrng, or the corresponding path of the TopSpin installation directory.

3. For ^{19}F data, open a dataset either with or without the investigated target.

4. To load the integration-file into the current spectrum, open the **Analyse** tab again, go to **Integrate** and using the **Read/Import integration regions** button, load the corresponding integration-file. This will load any defined regions of that file into the current spectrum.

5. Save and return to find a list of all integrated regions in the **Integrals** tab. Copy this into a spreadsheet or any other tool used for the further analysis of the data.

6. Repeat this procedure for every mix, with and without target.

4. Data Management

1. For ease of use and productivities sake, have a uniform work-flow set up for the further analysis and storage of the acquired data. For both the ^{1}H and the ^{19}F screening, use a specifically designed spreadsheet for each.

NOTE: For the ^{1}H screening this was purely used for data management and to summarize each target while for the ^{19}F screening it used the in chapter 4.3 explained quotient to automatically label each fragment as hit/no hit after the integral data was copied into it. This reduces the risk of human error during the analysis, assuming the file was set up properly, and makes sharing of information easier, as all the important information is gathered in one place in a file that can be opened by virtually anyone without the need of further programs for taking an initial look at the data.

Representative Results

Quality control of fragment library

The fragments from the in-house library were delivered as 50 mM stock solutions in 90% d6-DMSO and 10% D$_2$O (10% of D$_2$O ensures minimization of compound degradation due to repeated freeze-thaw cycles[14]). Single compound samples consisted of 1 mM ligand in 50 mM phosphate buffer (25 mM KPi pH 6.2 + 50 mM KCl + 5 mM MgCl$_2$), pH 6.0 in 90% H$_2$O/9% D$_2$O/1% d6-DMSO. ^{1}H-NMR experiments of fragments from the iNEXT library were measured on a 500/600 MHz NMR spectrometer. This data was further used for identifying the single compounds in ^{1}H screening campaigns using the CMC-q software which allows the user to fully acquire spectra in an automated manner and the analysis addon CMC-a the quality (solubility and integrity) of fragments was assessed. The results from the automated analysis from CMC-a are shown as a graphical output similar to what is represented in **Figure 3**. The graphical output

shows a representation of a 96-well plate. A red colored circle means that this fragment shows inconsistency in structure or concentration. Green colored wells indicate that the fragment is consistent.
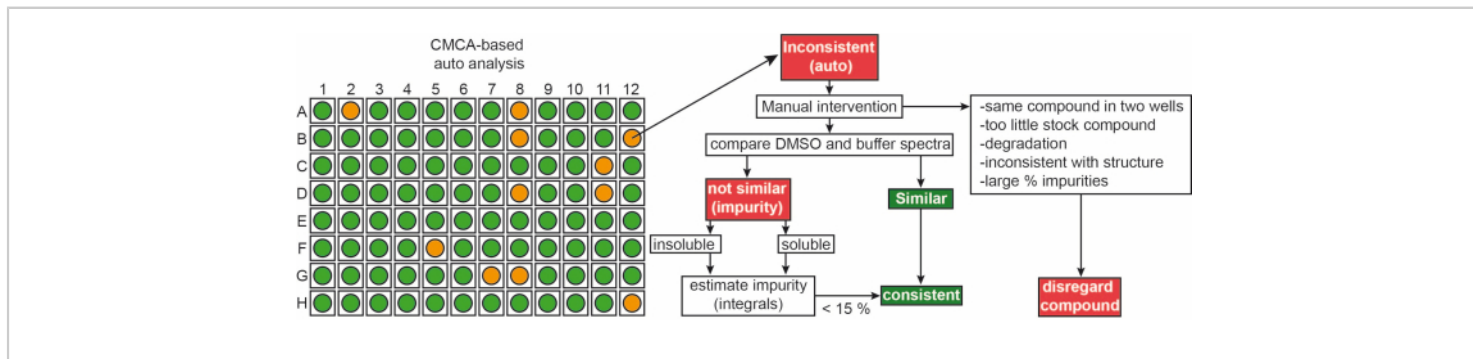

Figure 3: Quality control of fragment library. Schematic representation of CMC-a based automated output. Fragment properties such as concentration and structural integrity are assessed. Green stands for consistent, orange in this case stands for inconsistent. Inconsistent fragments are revised manually following the shown workflow. Please click here to view a larger version of this figure.

Approximately, 65% and 35% of the fragments were classified as consistent and in-consistent, respectively, in both DMSO and buffer. Further, 30% of the inconsistent classified ligands turned consistent after a careful manual inspection of the spectra[9].

^{19}F Mixture design

103 fragments containing one or several fluorine groups from the in-house library were divided into 5 mixes (A, B, C, D, E).

Each mix has 20 to 21 fragments. In this case the mixtures had to be carefully designed to avoid signal overlap. ^{19}F transverse relaxation experiments were measured for each mixture that apply CPMG pulse trains. These experiments can be modified by varying the relaxation delays. The ^{19}F chemical shift of mixes A-E can be seen in **Figure 4**.
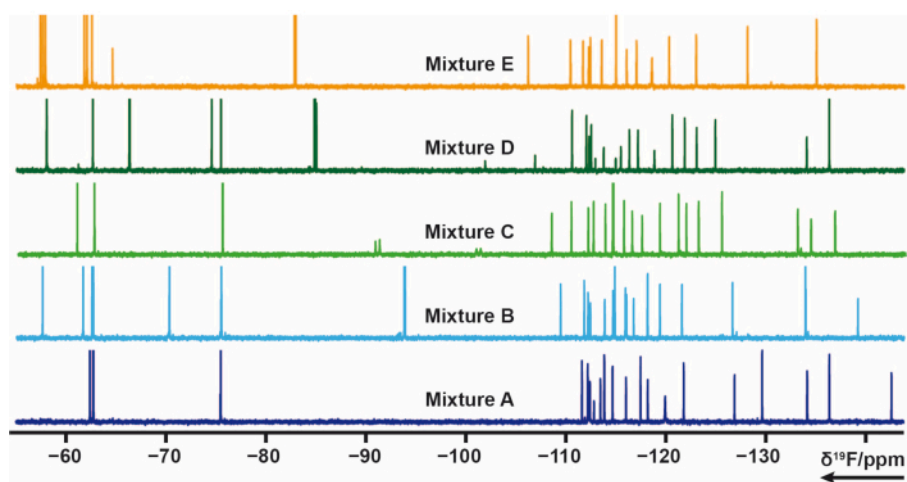
Figure 4: ^{19}F 1D-NMR spectra of mixture samples from the in-house library. Please click here to view a larger version of this figure.

Sample preparation

The sample preparation in the ^{19}F screening procedure was either done manually or with automated pipetting using a pipetting robot. The fragments in each mixture had a concentration of 2.5 mM in 90% d6-DMSO and 10% D_2O. The final volume of a screening sample was 170 µL with 5% D_2O as a locking agent. Each mixture was pipetted two times, one in a buffer containing solution (without target) and one into a target containing buffer solution. The ratio of target and fragment was set to 1:1, resulting in a final target/ligand concentration of 50 µM. Additionally control samples are the target biomolecule in screening buffer without a mixture to ensure target integrity as well as a control sample with only buffer and D_2O to ensure buffer quality.

NMR screening data of ^{19}F-1D and ^{19}F-CPMG-T_2 were measurements as described in section 3.1. For example, in the case of RNA a jump-return echo sequence (pp = zggpjrse,[15]) was acquired for the single target sample in buffer.

Data Analysis

The ^{19}F screening procedure was applied to the TPP riboswitch thiM from *E. coli* and protein tyrosine kinase (PtkA) from *M. tuberculosis* among several other targets[16]. The ^{19}F screening library has 103 fragments that are divided into 5 Mixes labelled from Mix A to E. Preparation of screening samples can be performed manually without the use of a sample pipetting robot. 40 µM thiM RNA containing solution (buffer conditions) was mixed with 3.2 µL from the mixtures. Further control samples were prepared consisting of buffer only, buffer with 5% of DMSO (previously ensure the stability of the biomacromolecule in the presence of the desired DMSO concentration) and buffer with RNA. These 13 screening samples were prepared and transferred to 3 mm NMR-tubes. Barcodes of NMR tubes are scanned and each mixture in the presence and absence of RNA, as well as control samples were measured according to the aforementioned ^{19}F NMR experiments performed at 298 K. Screening of thiM RNA against the in-house library was performed by conducting T_2 measurements with CPMGs of

0 ms and 200 ms for each different sample. Proper shimming and water suppression were monitored after finishing the measurements by comparing all DMSO peaks in terms of line broadening and intensity loss of additionally measured [1]H 1D experiments for all samples. Processing of obtained CPMG T_2 ^{19}F relaxation spectra was performed using a previously prepared and automated macro in TopSpin, respectively. Data analysis was performed following the instructions in the protocol section. The integral data obtained from

TopSpin (following the instructions in the protocol) can be evaluated quickly and easily using a pre-made spreadsheet or any similar program, by setting the correct conditions and thresholds. As described previously, thresholds are useful in defining binder, weak binder, or non-binder. **Figure 5** shows typical results of CPMG spectra of thiM RNA and PtkA, respectively. In some cases, further expert revision was needed.
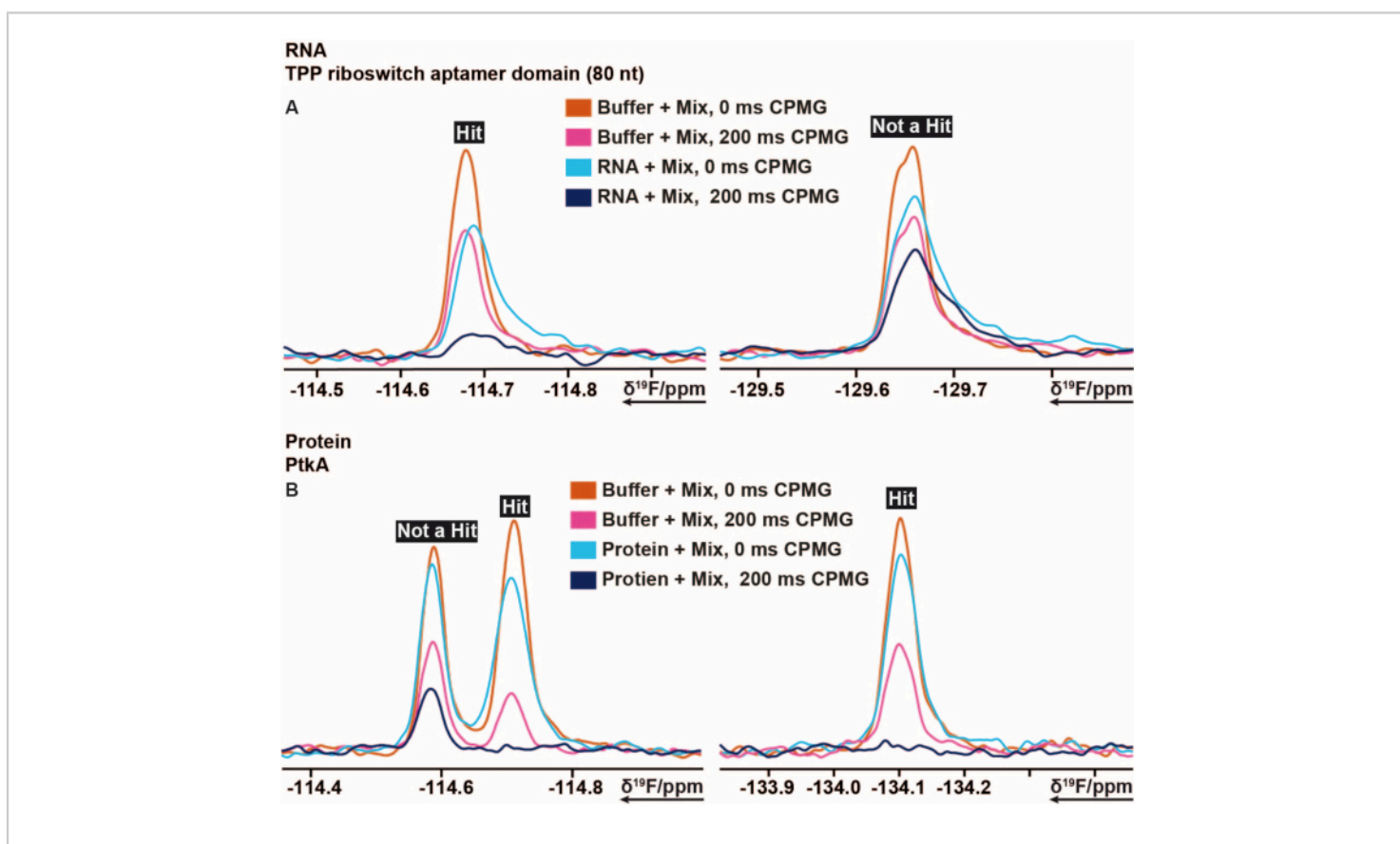
Figure 5: Cut out of ^{19}F CPMG NMR spectra showing the intensity changes obtained from different delay times of CPMG based experiments. (A) Representation of a binder (hit) and a non-binder in ^{19}F fragment-based screening performed on TPP riboswitch thiM RNA from *E. coli*. (B) Representation of a binder and a non-binder in ^{19}F fragment-based screening performed on PtkA from *M. tuberculosis*. Please click here to view a larger version of this figure.

[1]H Screening

Mixture design

The used in-house library is so diverse that for ^{1}H screening purposes no mixture design was performed. This means that 64 mixes were prepared by randomly choosing 12 to be mixing in one mixture.

Sample Preparation

For the ^{1}H screening of an exemplary SARS-CoV-2 RNA, automated pipetting using a pipetting robot was performed to prepare the samples. The fragments in each mixture had a concentration of 4.2 mM in 90% d6-DMSO and 10% D_2O. The final volume of a screening sample was 200 µL with 5% D_2O as a locking agent. 64 samples each containing a different mixture in 25 mM KPi, 50 mM KCl at pH 6.2 were pipetted without target RNA. Respectively, 64 samples were pipetted with target RNA, each containing a different mixture.

The RNA:Ligand ratio was set to 1:20, resulting in an RNA concentration of 10 µM and a ligand concentration of 200 µM.

Data Analysis

For the ^{1}H analysis, the FBS tool in TopSpin was used. To determine if a fragment is a hit, 1D chemical shift, waterLOGSY, and T_2 relaxation experiments were conducted. For T_2 relaxation, a decrease in intensity greater than 30% was counted as a hit, while for the chemical shift a shift of greater than 6 Hz was the cut-off. The waterLOGSY had to show a significant signal change (from negative to positive in this case). If any two of these three criteria were positive, a fragment was counted as a hit. Two examples for this can be seen in **Figure 6**.
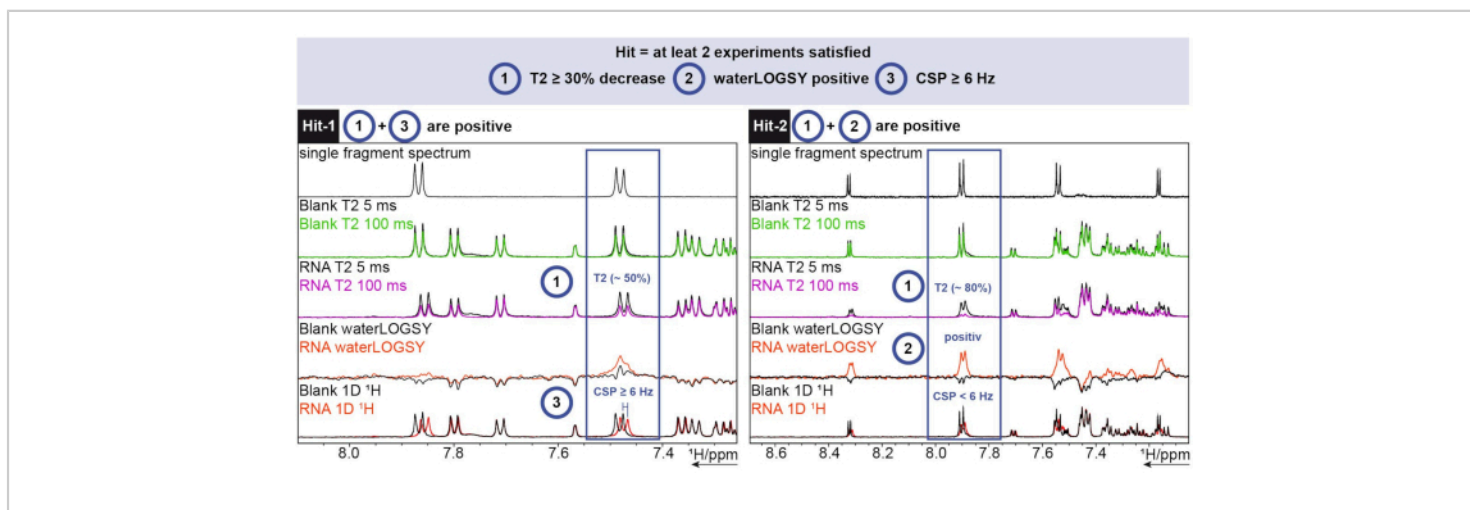
Figure 6: ^{1}H screening performed on an exemplary SARS-CoV-2 RNA showing hit determination criteria. Acquisition of three different experiments (^{1}H T_2 CPMG (5/100 ms), waterLOGSY, and 1D ^{1}H). Please click here to view a larger version of this figure.

Hit-1 shows a T_2 decrease of ~50% and a CSP ≥ 6 Hz. The waterLOGSY does not show a significant enough change in signal to also be counted as positive. As two out of three experiments are positive, this fragment is counted as a hit. For Hit-2, the T_2 shows a decrease of ~80% signal intensity and a clear signal change can be seen for the waterLOGSY. The CSP is not enough in this case, but as the two previous criteria are positive it is still counted as a hit.

Discussion

Versatility of the NMR-based fragment/drug screening. BMRZ has successfully implemented state-of-the-art automated NMR instrumentation as well as STD-NMR, waterLOGSY and relaxation experiments to identify fragments within a wide range of affinity regime for drug discovery. The installed hardware includes a high-throughput sample preparation robot and high-throughput sample storage, changer and data acquisition unit associated to a 600 MHz spectrometer. A recently purchased cryogenic probe for ^{1}H, ^{19}F, ^{13}C and ^{15}N ensures the required sensitivity for the proposed measurements and allows ^{1}H (1) decoupling during ^{19}F detection. This probe is connected to the latest generation of NMR console that offers the possibility to use the advanced software tools from Bruker, including CMC-q, CMC-assist, CMC-se and FBS (included in TopSpin). The fragment-based screening (FBS) tool is included in the latest version of TopSpin and helps to analyse the high-throughput data comprising of STD, waterLOGSY, T_2/T_{1r}-relaxation experiments. The liquid 1D 1H sample collection can be filled into the NMR-tubes in an automated manner by using the sample filling robot. Typically, a block of 96 tubes (3 mm) are filled in approximately two hours. The 96-well-plate-racks are directly positioned in the HT sample changer, which reads the barcode of the block and assigns the NMR tubes to the experiments controlled by the automation software (*IconNMR*). Five 96-well-plate-racks can be stored and programmed in the HT sample changer at the same time. The temperature of each of the individual racks can be controlled and regulated separately. Additionally, each individual sample can be preconditioned (preheating and tube drying for removal of condensed humidity) to the desired temperature before the measurement.

Suitability for wide range of applications. One of the broad applications of this automated NMR-based screening is to identify and develop novel ligands binding to a biomacromolecular target (DNA/RNA/Proteins). These ligands can include orthosteric and allosteric inhibitors that typically bind non-covalently. Further, FBDD by NMR is typically used as a first step to select promising compounds, the requirements to be met are availability of the biomolecular target in sufficient quantities. This objective is divided into two major tasks.

Task one is to develop and characterize an in-house fragment library for the following reasons: initial and periodic quality control, characterization, and quantification of more than 1000 fragments; determination of solubility of the fragments in buffers optimized for each target, in particular for protein targets; and the establishment of several libraries to accommodate diverse scaffolds and extending towards other macromolecule classes. Task two is to integrate workflows for fragment-based drug design (FBDD) by NMR using: automated 1D-ligand observed screening (^{1}H and ^{19}F observed); automated replacement assays (competition experiments with (natural) ligand) to differentiate orthosteric and allosteric binding; automated secondary screenings with multiple fragments; automated 2D-protein screening, and secondary screening of a set of derivatives around an initial hit making use of the EU-OPENSCREEN library or any other library; and re-profiling screening of FDA-library against the chosen targets.

Additionally, metabotyping of various cell lines (disease relevant) can be conducted in order to unravel the regulatory mechanisms that link cell cycle control and metabolism. Also, there is functional characterization of RNA/DNA/protein regulation elements *in vivo* and *in*

vitro for optimization of construct/domain optimization (stability optimization for structural investigations (Buffer, pH, temperature, and salt screening), and an extension of NMR-based fragment screening to membrane proteins and intrinsically disordered proteins, which are generally inaccessible to other techniques.

Limitations. Use of ^{19}F and ^{1}H fragments libraries have their pros and cons, few of which will be mentioned in the following. The largest benefit of ^{19}F versus ^{1}H measurements is the speed of both the actual measuring time and the subsequent analysis, as the mixtures contain almost double the number of fragments and fewer experiments must be conducted. The follow up analysis is also easier for ^{19}F screening, as there is no interference from buffers and additionally offers a broader chemical shift range with almost no signal overlap for an optimally designed fragment mixture. The spectra themselves are greatly simplified, usually only having one or two signals per fragment, depending on the number of fluorine atoms. The analysis of these spectra can therefore be automated, again cutting down on time. This comes at the cost of chemical diversity, at least for the library used in this study. As only ~13% of the library contains ^{19}F, but naturally all of them are useable in ^{1}H screening, the diversity of the ^{19}F screening fragments will be lower. This could be circumvented using specifically designed ^{19}F libraries with more fragments and bigger chemical diversity. Another disadvantage for ^{19}F screening is the low number of signals per fragment. Fragments generally are composed of more than one hydrogen atom. Therefore, ^{1}H observed screening experiments can rely on different signals for the same fragment for detecting binding. This gives a higher degree of confidence when identifying hits for the ^{1}H screening,

whereas the ^{19}F screening must rely on the one or two signals given per fragment.

A detailed account on the modern automated NMR-based fragment screening instrumentation, software and analysis methods and protocols thereof has been presented. The installed hardware includes a high-throughput sample preparation robot and a high-throughput sample storage, changer and data acquisition unit associated to a 600 MHz spectrometer. A recently installed cryogenic probe head for ^{1}H, ^{19}F, ^{13}C and ^{15}N ensures the required sensitivity for the proposed measurements and allows ^{1}H decoupling during ^{19}F detection. Further, the latest generation of NMR console offers the possibility to use advanced analytical software for aiding acquisition and on-the-fly analysis. The above discussed technology, workflows, and the described protocols should foster remarkable success to users pursuing FBS by NMR.

Disclosures

None.

Acknowledgments

References

1. Yanamala, N. et al. NMR-Based Screening of Membrane Protein Ligands. *Chemical Biology & Drug Design*. **75**, 237-256 (2010).

2. Souers, A. J. et al. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nature Medicine.* **19**, 202-208 (2013).

3. Su, M. C., Chang, C. Te, Chu, C. H., Tsai, C. H., Chang, K. Y. An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. *Nucleic Acids Research.* **33**, 4265-4275 (2005).

4. Perera, T. P. S. et al. Discovery & pharmacological characterization of JNJ-42756493 (Erdafitinib), a functionally selective small-molecule FGFR family inhibitor. *Molecular Cancer Therapeutics.* **16**, 1010-1020 (2017).

5. Zhang, C. et al. Design and pharmacology of a highly specific dual FMS and KIT kinase inhibitor. *Proceedings of the National Academy of Sciences of the United States of America.* **110**, 5689-5694 (2013).

6. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews.* **23**, 3-25 (1997).

7. Congreve, M., Carr, R., Murray, C., Jhoti, H. A 'Rule of Three' for fragment-based lead discovery? *Drug Discovery Today.* **8**, 876-877 (2003).

8. Chávez-Hernández, A. L., Sánchez-Cruz, N., Medina-Franco, J. L. A Fragment Library of Natural Products and its Comparative Chemoinformatic Characterization. *Molecular Informatics.* **39**, 2000050 (2020).

9. Sreeramulu, S. et al. NMR quality control of fragment libraries for screening. *Journal of Biomolecular NMR.* doi:10.1007/s10858-020-00327-9 (2020).

10. Gao, J. et al. Automated NMR Fragment Based Screening Identified a Novel Interface Blocker to the LARG/RhoA Complex. *PLoS One.* **9**, e88098 (2014).

11. Peng, C. et al. Fast and Efficient Fragment-Based Lead Generation by Fully Automated Processing and Analysis of Ligand-Observed NMR Binding Data. *Journal of Medicinal Chemistry.* **59**, 3303-3310 (2016).

12. Cox, O. B. et al. A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chemical Science.* **7**, 2322-2330 (2016).

13. Hwang, T. L., Shaka, A. J. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *Journal of Magnetic Resonance, Series A.* **112**, 275-279 (1995).

14. Gossert, A. D., Jahnke, W. NMR in drug discovery: A practical guide to identification and validation of ligands interacting with biological macromolecules. *Progress in Nuclear Magnetic Resonance Spectroscopy.* **97**, 82-125 (2016).

15. Sklenar, V., Bax, A. A new water suppression technique for generating pure-phase spectra with equal excitation over a wide bandwidth. *Journal of Magnetic Resonance.* **75**, 378-383 (1987).

16. Binas, O. et al. ^{19}F NMR-Based Fragment Screening for 14 Different Biologically Active RNAs and 10 DNA and Protein Counter-Screens. *ChemBioChem.* cbic.202000476 (2020).