DOI: 10.1070/RCR4836

NMR screening and studies of target – ligand interactions

Vladimir I. Polshakov, Evgeniy A. Batuev, Alexey B. Mantsyzov

Faculty of Fundamental Medicine, Lomonosov Moscow State University
Lomonosovsky prosp. 27, stroenie 1, 119991 Moscow, Russian Federation

Nuclear magnetic resonance (NMR) spectroscopy is one of the leading biophysical methods used in the search for and design of physiologically active compounds considered as potential drugs. The review concerns modern NMR techniques used to study the binding of low-molecular-mass compounds to biomacromolecular targets. The most promising methods of NMR screening and strategies for rational lead design are discussed. They were used to design drugs that have been approved for the use in medical practice or are in the final stages of clinical trials. Examples are given of the application of the fragment-based drug design and NMR screening techniques to the design of novel drugs.
The bibliography includes 252 references.

Contents

1. Introduction 59
2. Interaction of small molecules with target proteins 61
3. NMR spectroscopy methods to study protein – ligand interactions 63
4. NMR screening methods 73
5. NMR spectroscopy methods for lead design 83
6. Examples of application of fragment-based approach and NMR methods to the design of novel drugs 88
7. Conclusion 91

1. Introduction

In the early 20th century, rapid progress in natural sciences and the development of novel methods of synthesis and structure determination of organic compounds led to a situation where the search for novel drugs got a strong chemical basis. Since then drug design has involved the synthesis of novel structural analogues of metabolites, endogenous compounds, hormones, *etc.* rather than simple application of the empirical method known since ancient times. A hundred years later the science of novel drug design is based on biology rather than chemistry. To date, key biological targets have been established, methods for structure determination of biomolecules (X-ray crystallography, nuclear magnetic resonance, cryoelectron microscopy) developed, genetic engineering methods for production, isolation and purification of designed biomacromolecules, as well as methods of investigation of interactions between the targets and potential drugs. Target-oriented methods of search for novel lead compounds have become a basis for modern drug design strategies.[1, 2]

Key factors simulating research activity in this field include expansion of the list of diseases, drug resistance and identification of new side effects of the known drugs. The development of a novel drug is a multistep process that requires huge, constantly growing time demand and cost (Fig. 1).† Leading pharmaceutical companies estimate that the development of an innovative drug may take 10 to 15 years and >1 billion US dollars.[3] The overall cost is mainly spent for long-term and expensive phases of clinical trials of drugs, where cost reduction is impossible since the requirements for the safety and efficacy of the drugs being designed are continuously toughened.[4, 5] Up to 25% of time and cost are spent for the initial steps of hit discovery and hit to lead optimization. These expenses can be reduced by using modern

V.I.Polshakov. PhD, DSc, Leading Researcher of the Laboratory of Magnetic Tomography and Spectroscopy at the Faculty of Fundamental Medicine, MSU.
Telephone: +7 (495) 939 – 4485, e-mail: vpolsha@mail.ru
Current research interests: NMR spectroscopy, protein structure and dynamics, protein – ligand interactions, molecular pharmacology.
E.A.Batuev. Researcher of the same Laboratory and Young Researcher of the Research Institute of Applied Acoustics.
Telephone: +7 (495) 939 – 4485, e-mail: carboksyl@ya.ru
Current research interests: NMR spectroscopy, structure of organic compounds and peptides, NMR screening.
A.B.Mantsyzov. PhD, Researcher of the Laboratory of Magnetic Tomography and Spectroscopy at the Faculty of Fundamental Medicine, MSU.
Telephone: +7 (495) 939 – 4485, e-mail: mants@inbox.ru
Current research interests: NMR spectroscopy, protein structure and dynamics, molecular dynamics, docking, NMR screening.

Received 22 May 2018
Translation: A.M.Raevskiy

† All figures were compiled by the authors using materials reported in the references cited (see particular Figure Captions).

© 2019 Uspekhi Khimii, Russian Academy of Sciences and Turpion Ltd.

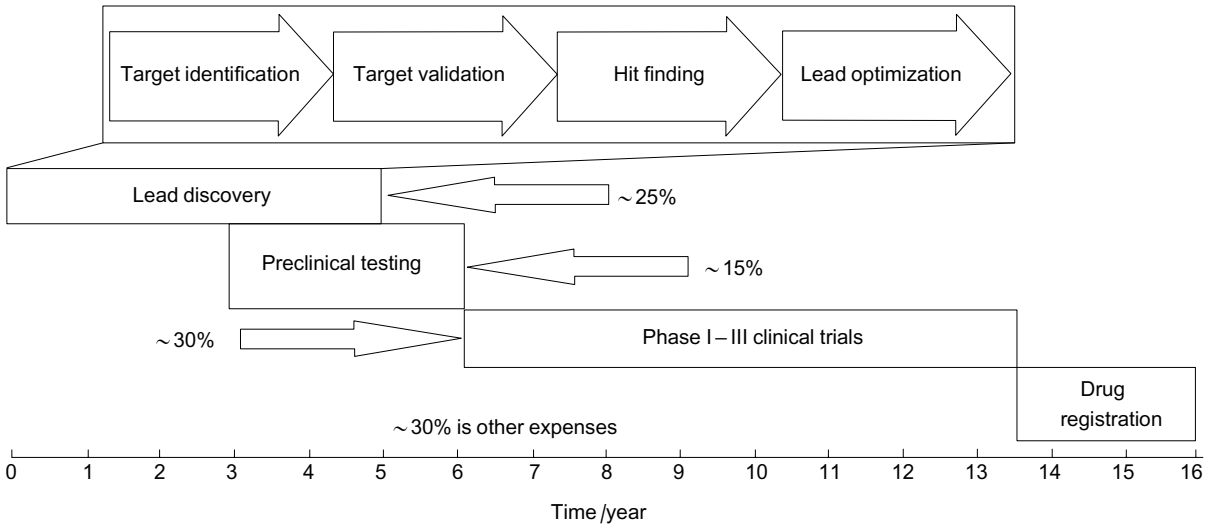


Figure 1. Average periods and expenses for main steps in the drug discovery process (in percentage terms).[4,5] Upper panel: key stages of target-oriented search for a novel physiologically active compound — drug candidate.

strategies of rational design of physiologically active compounds based on information on biological targets.

Identification and validation of a new target are carried out using methods of molecular biology,[1] bioinformatics,[6] genomics[7] and proteomics.[8] In recent years, high-tech methods of phenotypic screening on cell models, protozoa, or insects have been rapidly developed, leading to the discovery of new biological targets.[9,10] Targets can include enzymes, transport proteins, receptors, nucleic acids, as well as nucleoprotein complexes.[11,12]

Currently, a few hundred targets are involved in the search for new drugs; a few thousands of potential targets were identified.[11] Having identified and characterized a target, one can use target-oriented methods to develop novel biologically active compounds (hits).[13] Such methods are based on the design of chemical compounds that can rather strongly bind to targets and thus change their functional properties. Note that there is no need to find compounds with extremely high affinity for targets. The dissociation constants of the active substances of drugs currently used in medical practice from complexes with respective targets vary from 10^{-4} to 10^{-10} M with a maximum occurrence of about $\sim 10^{-8}$ M.[14]

Undoubtedly, there is vast gap between identification of a hit compound capable of binding a target and registration of a drug. The new hit compound can exhibit toxic properties (this should be revealed during preclinical testing and clinical trials), demonstrate poor absorption and extremely rapid excretion, undergo rapid metabolism, fail to penetrate blood–tissue barriers in the organism, *etc*. As a result, this compound can simply not reach its pharmacological targets. Lipinski *et al*.[15] analyzed the structure and physicochemical profiles of most active substances used in medical practice and devised a rule called the 'Rule of 5'. An orally administered active substance obeying this rule is expected to have high bioavailability and a favourable pharmacokinetic profile. More exactly, Lipinski's 'Rule of 5' for orally administered active substances assumes at most one violation of the following properties: the molecule should

— contain not more than five hydrogen bond donors (NH or OH groups);

— contain not more than ten hydrogen bond acceptors (N, O or S atoms);

— have a molecular mass of 500 Da or less; and

— have a calculated octanol–water partition coefficient (log*P*) of at most 5.

Compounds meeting these criteria are thought to be druglike. More recently, numerous extended versions of the 'Rule of 5' were reported, which include new criteria, *e.g*., molecular surface area of a compound, number of rotatable bonds, *etc*.[16,17] At present, screening or design of new compounds is almost always carried out taking account of these empirical rules.

Among experimental methods of investigation used in the field of drug desing, nuclear magnetic resonance spectroscopy holds a prominent place. In this review we will discuss the application of NMR spectroscopy methods in studies of protein–ligand interactions and modern NMR screening methods aimed at the design of physiologically active compounds.

There is considerable recent literature on NMR screening methods and NMR-based design of lead compounds. Pioneering reviews on the title topic date back to the years 1999–2004.[18–26] They were followed by reviews on NMR spectroscopy methods for quantification of protein–ligand binding,[27–29] application of the fragment-based methods and NMR screening for the design of high-affinity ligands,[2,30–36] comparison of the potential of NMR spectroscopy and other biophysical methods in the search for structural fragments.[37,38] Yet another review[2] covers a two-decade evolution of fragment-based drug discovery and presents a number of examples.

This review is devoted to fundamental aspects of protein–ligand interactions that are crucial for appropriate consideration of methods for detection of specific ligand–target binding. We will analyze the capabilities of

NMR spectroscopy for monitoring protein – ligand interactions and NMR screening. In Section 7 we will discuss the lead design strategies and present examples of either approved compounds or those being in Phase II/III clinical trials.

Notations are as follows:

ALARM-NMR — A La Assay to detect reactive molecules by nuclear magnetic resonance;

BACE1 — β-secretase 1;

Bcl-2 — intracellular protein factor, regulator of programmed cell death;

CDK2 — cyclin-dependent kinase 2;

CPMG — Carr – Purcell – Meiboom – Gill pulse sequence;

CRINEPT — cross-correlated relaxation-enhanced polarization;

CSF1 — colony stimulating factor 1;

DHFR — dihydrofolate reductase;

DOSY — diffusion-ordered spectroscopy;

DSF — differential scanning fluorimetry;

DTT — dithiothreitol;

FAXS — fluorine chemical shift anisotropy and exchange for screening;

n-FABS — n-fluorine atoms for biochemical screening (method is analogous to FAXS except that it uses a number of chemically equivalent ^{19}F nuclei);

FBDD — fragment-based drug design;

FBLD — fragment-based lead design;

GE — group efficiency;

HSQC — heteronuclear single quantum correlation;

HTS — high-throughput screening;

INPHARMA — interligand NOE for pharmacophore mapping;

ITC — isothermal titration calorimetry;

NADP — nicotinamide adenine dinucleotide phosphate;

NADPH — reduced form of NADP;

NMR — nuclear magnetic resonance;

NMR-DOC — nuclear magnetic resonance docking of compounds;

NMR-SOLVE — NMR structurally oriented library valency engineering;

NOE — nuclear Overhauser effect;

NOESY — nuclear Overhauser effect spectroscopy;

RAMPED-UP NMR — rapid analysis and multiplexing of experimentally discriminated uniquely labeled proteins using NMR;

ROESY — rotational nuclear Overhauser effect spectroscopy;

PPAR — peroxisome proliferator-activated receptor;

PrOF NMR — 1D-protein-observed fluorine NMR;

SALMON — solvent accessibility, ligand binding, and mapping of ligand orientation by NMR spectroscopy;

SAR by ILOE — structure – activity relationships by interligand Overhauser effect;

SAR by NMR — structure – activity relationships by NMR spectroscopy;

SHAPES — a NMR screening strategy based on the use of a small but diverse library of ligand structures;

SLAPSTIC — spin labels attached to protein side chains as a tool to identify interacting compounds;

SPR — surface plasmon resonance;

STD — saturation transfer difference;

TINS — target immobilized NMR screening;

TMP — trimethoprim;

TOCSY — total correlation spectroscopy;

trNOE — transferred NOE;

TROSY — transverse optimized correlation spectroscopy;

WaterLOGSY — water-ligand observed *via* gradient spectroscopy.

Notations of main physical quantities:

k_{on} — rate constant for ligand – protein binding;

k_{off} — rate constant for ligand dissociation from complex;

K_d — dissociation constant;

K_a — binding constant;

ΔG — change in Gibbs free energy;

ΔH — change in enthalpy;

ΔS — change in entropy;

LE — ligand efficiency;

ν — frequency of a signal;

δ — chemical shift;

R_2 — transverse relaxation rate;

T_1 — longitudinal relaxation time;

T_2 — transverse relaxation time;

I — signal intensity;

τ — time delay;

τ_c — correlation time of rotational diffusion of a molecule;

$\delta\omega$ — resonance frequency difference;

$R_{1\rho}^{obs}$ — relaxation rate in the rotating frame;

υ — proportion of ligand;

γ — gyromagnetic ratio;

t_m — mixing time;

IC_{50} — inhibition constant.

2. Interaction of small molecules with target proteins

2.1. Basic thermodynamic and kinetic aspects of the interaction of low-molecular-mass ligands with proteins

Prior to discussing methods of search for compounds capable of binding the target proteins and thus changing their functional properties we will consider key factors determining the protein – ligand binding processes.

In the simplest case (one binding site) the interaction of a ligand with the target protein is described by the equilibrium between the free forms of the ligand (L) and protein (P) and the protein – ligand complex being formed (PL)

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL \qquad (1)$$

The rate constant for protein – ligand binding (ligand association rate), k_{on}, depends primarily on the ligand diffusion rate and lies in the range of $(1-3)\times10^7$ M^{-1} s^{-1}, although it can vary significantly depending on the size of the ligand and on the electrostatic properties of the ligand and protein.[39] The rate constant for ligand dissociation from complex, k_{off}, is determined by the strength of protein – ligand interactions.

The dissociation constant K_d and its reciprocal, the binding constant K_a, are equal to the k_{off}/k_{on} ratio

$$K_d = K_a^{-1} = \frac{[P][L]}{[PL]} = \frac{k_{off}}{k_{on}} \quad (2)$$

The K_d constant is related to the change in the Gibbs free energy upon formation of the protein–ligand complex:

$$\Delta G = (G_P^{free} + G_L^{free}) - G_{PL}^{complex} = -RT \ln K_d \quad (3)$$

The calculated values of the parameters K_d, k_{off}, ΔG and the proportions of the bound form of the ligand corresponding to formation of protein–ligand complexes of different strength are listed in Table 1.

The total change in the Gibbs free energy depends on the enthalpy (ΔH) and entropy ($T\Delta S$) contributions:

$$\Delta G = \Delta H - T\Delta S = \Delta H - T(S_P^{free} + S_L^{free} - S_{PL}^{complex}) \quad (4)$$

The change in enthalpy, ΔH, is mainly due to interatomic ligand–protein interactions. Ligand-induced structural rearrangements in the protein molecule may also appear to be significant.[40] Typical contributions of key interatomic interactions to the free energy of protein–ligand binding are listed in Table 2.[41–45] More exactly, these interactions primarily contribute to the change in enthalpy.

Hydrophobic interactions are determined by the changes in the entropy of the system containing a large number of water molecules. Nevertheless, hydrophobic ligand–protein interactions contribute to the binding free energy ΔG similarly to Coulomb interactions or hydrogen bonding. The contributions to the change in the Gibbs free energy upon transfer of certain hydrophobic groups from aminoacyl-tRNA synthetase protein to water are as follows:[46]

Table 1. Calculated values of parameters K_d, k_{off} and ΔG for the formation of protein–ligand complex ($T = 298$ K) and the proportion of the bound form of the ligand assuming equimolar initial protein/ligand ratio and concentrations of 1 mM for both protein and ligand.

K_d /M	k_{off} /s^{-1}	ΔG /kJ mol^{-1}	Proportion of the bound form of ligand
		Weak binding	
10^{-2}	10^5	11.4	0.084
10^{-3}	10^4	17.1	0.382
10^{-4}	10^3	22.80	0.730
		Moderately strong binding	
10^{-5}	10^2	28.5	0.905
10^{-6}	10	34.2	0.969
10^{-7}	1	39.9	0.990
		Strong binding	
10^{-8}	10^{-1}	45.6	0.997
10^{-9}	10^{-2}	51.3	0.999
10^{-10}	10^{-3}	57.1	~ 1

Note. All values were calculated using Eqns (2), (3) and (13) (see below); the k_{off} values were determined for $k_{on} = 10^{-7}$ mol^{-1} s^{-1}.

Table 2. Contributions of interatomic interactions to the free energy of protein–ligand binding.

Type of interaction (contribution)	Energy /kJ mol^{-1}	Ref.
Hydrogen bond	11.7–37.6	41, 43
Coulomb interaction	0.8–4.2	41
π-Stacking interaction	2.0–6.3	44, 45
Cation-π interaction	15.9–71	42
Dipole–dipole (orientation) interaction	0.8–2.0	41
Debye (induction) interaction	0.08–2.0	41
London (dispersion) interaction [a]	< 0.008	41

[a] The energies of particular dispersion interactions are low; however, the total energy of dispersion interactions in the entire molecule can be rather high and span the range of 1 to ~50 kJ mol^{-1}.[41]

Group	Me	Et	Pr^i
Energy /kJ mol^{-1}	13.8	27.2	40.2

The change in the entropy of the system on going from free ligand and protein to their complex should also be taken into account.[41,47] In the case in point, it leads to a decrease in the free energy of protein–ligand interaction and consequently in decrease of K_d.

The entropy (TS) contributions of translational (I) and rotational (II) motions of a small ligand molecule (~200 Da), rotation about a single bond (III) and intramolecular vibrations (IV) to the Gibbs free energy (at 298 K)[41,47] are given below.

Type of motion	*I*	*II*	*III*	*IV*
TS /kJ mol^{-1}	36–45	12 – 34	2–3	0.2–2

As can be seen, the major contributions come from the entropies of translational and rotational motions, which decrease considerably upon complexation. It is these effects that are responsible for well-known enthalpy-entropy compensation of the total energy of protein–ligand interactions.[41] The formation of a relatively strong complex with K_d of the order of 10^{-7} M (*i.e.*, ΔG of the order of 39 kJ mol^{-1}, see Table 1) requires the formation of a network of intermolecular interactions with energies in the range of 90–120 kJ mol^{-1} in order to compensate the loss due to the change in the translational and rotational entropy of the ligand (50–79 kJ mol^{-1}). Most often, the entropy contribution appears to be even larger owing to molecular vibrations and rotation about single bonds. The entropy contribution from rotation of ligand fragments about single bonds can be evaluated using the Boltzmann formula

$$S = k \ln W, \quad (5)$$

where k is the Boltzmann constant and W is the number of microscopic states. For most single C–C bonds, the number of stable conformers equals three. As a result, the free energy of the ligand decreases by ~2 kJ mol^{-1} (for 298 K), which is comparable with the energy of Coulomb interaction. That is why rigid ligands are more

tightly bound to the target proteins than conformationally flexible ones, all other things being equal.

Bulkier ligands bearing a larger number of functional groups that can participate in the interatomic interactions with proteins always have higher binding energies compared to small molecules. To take into account the size of the ligand, the parameter 'ligand efficiency' (LE) was introduced. By default, LE equals the binding free energy normalized to the number of heavier (non-hydrogen) atoms of the ligand (N),[48]

$$\mathrm{LE} = \frac{-\Delta G}{N} \qquad (6)$$

Binding of a ligand with a molecular mass of the order of 500 Da to its target with $K_d \sim 10$ nM requires LE > 1.3 kJ mol^{-1}. More recently, the parameter 'group efficiency' (GE) was introduced[49] to take into account the contributions of particular functional groups to the total ligand binding energy; the GE values are calclated by analogy with LE ones.

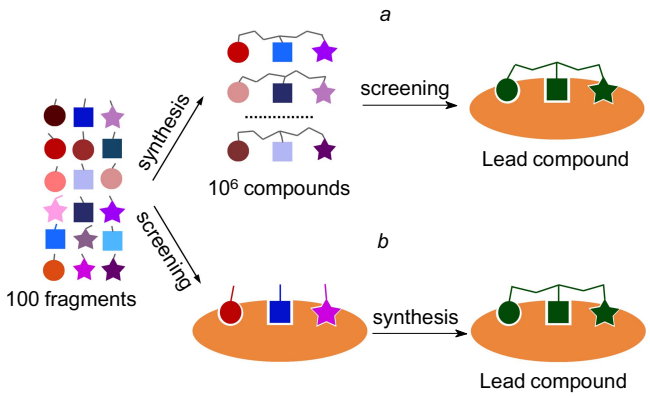


Figure 2. Approaches based on high-throughput screening (*a*) and fragment-based lead design (*b*): a comparison.[38]
In case (*a*) one should synthesize a much larger number of compounds for screening.

2.2. Fragment-based lead design approaches

Target-oriented lead-finding strategies involve various approaches. If the target structure is available, various molecular docking (*in silico*) methods are most often used, followed by experimental validation of the hits for interaction with the target.[50] High-throughput screening (HTS) of large ligand libraries is also widely used.[51] Recently, there has been rapid progress in fragment-based frug design (FBDD) and fragment-based lead design (FBLD).[52, 53]

Methods involving fragment-based screening, linking and final structure optimization of the ligand molecule have been widely used in recent years.[53–56] Practical implementation of these methods does not require knowledge of the target structure (it is desired but not necessary). In accordance with 'The Rule of Three' (analogue to Lipinski's 'Rule of Five', see above),[57, 58] molecular fragments should be small and usually have a mass from 100 to 300 Da, a limited number of donors and acceptors of hydrogen bonds (< 3) and sufficiently rigid structures where the number of freely rotating bonds is also at most three. To design a lead compound, one should identify, optimize and correctly link two or three structural fragments. This approach makes it possible to significantly increase the structural diversity of the compounds studied. The HTS libraries rarely exceed 10^6 compounds; at the same time, this number of compounds can be designed by combinatorial linking of three fragments chosen from the library that includes only 100 compounds (Fig. 2).

Structural fragments of a potential lead compound binds the target with relatively low affinity (dissociation constants from the complex are in the range of 10^{-4} to 10^{-2} M). However, linking of even two fragments leads to much more efficient binding since the translational and rotational entropy of the newly formed molecule differ only slightly from the entropy of a single fragment. Thus an additional contribution to the binding free energy appears due to the absence of undesired entropy contribution of the second fragment.[59] As mentioned above, such an entropy contribution can be as large as 50–80 kJ mol^{-1} for a fragment with a molecular mass ~ 200 Da; this corresponds to a decrease in the dissociation constant by a few orders of magnitude. It should be noted that a small molecular fragment can be comparable with a large molecule in the parameter LE.[60, 61] Therefore, fragment screening can be performed with the same criterion (LE > 1.3) as that used for large ligands (see above), although binding between small fragments and the target will be rather weak.

Only a few methods implemented in HTS protocols can be used to monitor weak protein–ligand interactions observed in the case of structural fragments. One should use methods that can distinguish between weak specific interactions in the binding site of a protein and nonspecific binding with, *e.g.*, charged amino acid residues on the surface of the target protein. Inability to discriminate specific and nonspecific binding leads to higher false-positive rates. Main physical methods used to search for structural fragments are characterized in Table 3. These methods are well documented (see Refs 62–69, 73–78); each of them has its advantages and drawbacks.

An analysis of the data presented in Fig. 3 allows one to compare the methods mentioned above in ability to detect protein–ligand interactions and provide structurally or functionally useful information provided that the dissociation constants of ligands vary from a few mM^{-1} to a few pM^{-1}. The frequency distribution of the binding constants of marketed drugs to the target proteins has a maximum at $K_a \approx 10^8$ M^{-1}. Besides, a large number of compounds can bind to the target proteins more or less tightly, the observed affinity differences being as high as eight orders of magnitude and even more.

3. NMR spectroscopy methods to study protein–ligand interactions

NMR spectroscopy is one of the most informative methods of investigation of ligand–target interactions. Owing to the contribution of R.Ernst and K.Wüthrich, 1991 and 2002 Nobel Prize laureates in chemistry, and many other researchers, this method has become a major tool in structural biology. A distinctive feature of NMR spec-

Table 3. Characterization of methods for detecting fragment – target interactions.

Method	Ligand concentration /mM	Required amount of protein /μg	Main advantages and drawbacks	Ref.
X-Ray crystallography	10^{-4} – 10	$(5-50)\times10^3$	Information on 3D structure; high-quality protein crystals are required, withstanding the action of concentrated ligand solutions	62 – 64
Surface plasmon resonance	10^{-6} – 10^2	1 – 10	Information on the binding and dissociation kinetics; immobilization of proteins is required; high false-positive rates	65
Biolayer interferometry	10^{-6} – 10^2	1 – 10	Information on the binding and dissociation kinetics; search for inhibitors of protein – protein interaction; high false-positive rates	66
Isothermal titration calorimetry	10^{-3} – 1	50 – 200	Information on thermodynamics of protein – ligand binding and dissociation constants; high protein concentrations are required; false-positive results are possible	67
Differential scanning fluorimetry	0.1 – 1	50 – 200	Identification of structure stabilization of ligand-bound proteins; not all proteins ca be studied by this method; false-positive results are possible	68
Mass spectrometry	10^{-5} – 1	< 1	High-sensitivity; no need for binding to a surface or isotope labelling of targets; problems related to aggregation of ligands; high false-positive rates	69, 70
Microscale thermophoresis	10^{-6} – 10^2	1 – 10	High sensitivity; the possibility of studying almost all types of ligands and macromolecules; fluorescent labels are required	71,72
NMR spectroscopy with detection of target resonances	10^{-3} – 10	$(5-50)\times10^3$	Information on the ligand binding site and on the structure of the complex; large amount of protein and protein labelling with stable isotopes are required	23
NMR spectroscopy with detection of ligand resonances	10^{-2} – 10	$50-5\times10^3$	Information on ligand – target binding	38

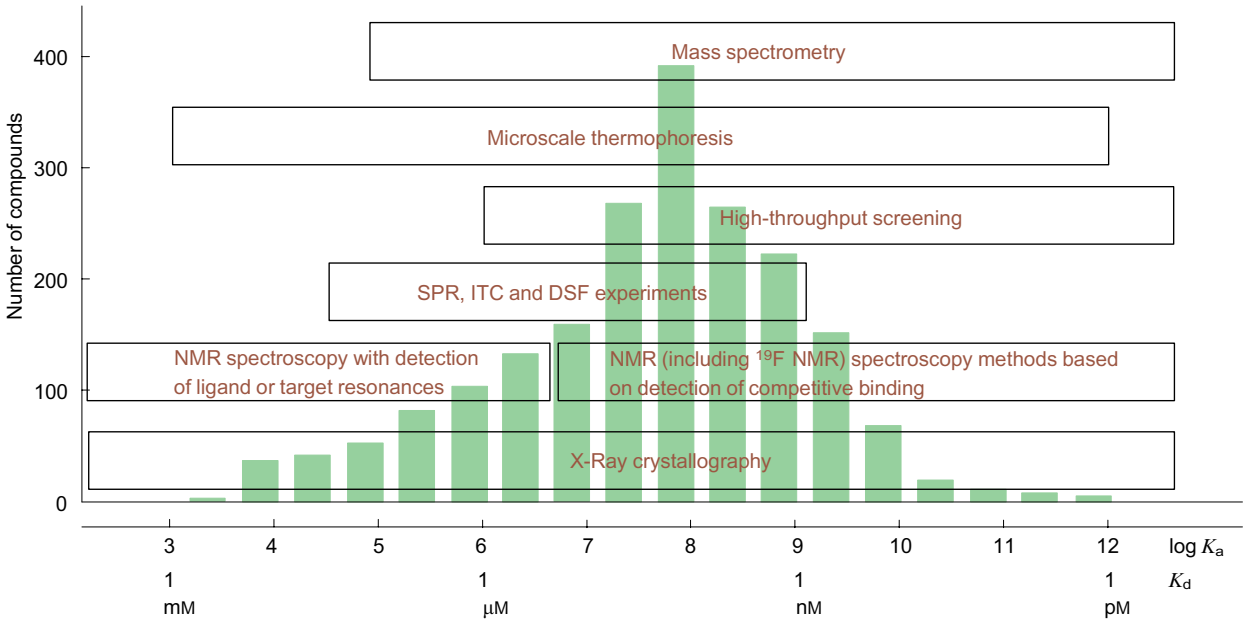


Figure 3. Main methods used to detect ligand binding to protein targets.[14]
Shown are the ranges of the binding constant values corresponding to efficient application of the respective methods. The frequency distribution of the binding constants of marketed drugs to target proteins (green rectangles) includes only the drugs with known K_a values and explicit target associations.

troscopy is the possibility to identify specific interactions between a ligand and the target protein even in the case of weak binding. Various NMR spectroscopy techniques allow one:

— to determine the structures of the target and ligands interacting with it;

— to establish the conformation of the bound ligand;

— to map the interface of the target – ligand interactions;

— to determine the ligand – target binding stoichiometry;

— to evaluate the ligand – target association constants;

— to measure the kinetics of ligand – target binding;

— to estimate the dynamic characteristics of the target and ligands binding to it;

— to search for binding ligands; and

— to determine the mutual arrangement of two and more ligands in the binding sites of the target.

The abundance of NMR data on protein – ligand interactions is based on the diversity of spectral parameters (Table 4). Undoubtedly, one of the most important NMR parameters is the chemical shift. Signals of a ligand or amino acid residues of a protein interacting with the ligand change upon the formation of the protein – ligand complex since the chemical environment of the detected nucleus changes.[73] Note that significant changes in the chemical shifts are observed only in the case of specific ligand – protein binding. Nonspecific interactions cause no marked changes in the chemical shifts because the binding effects are averaged owing to a wide variety of possible ligand arrangements relative to the protein.

The choice of the measured parameters and corresponding experimental NMR techniques depend on the character of the protein – ligand interaction and, first of all, on the rate of exchange between the free and bound forms. The NMR spectra of a ligand obtained for the case of equilibrium between the free and protein-bound forms, are presented in Fig. 4. When the rate constant for ligand dissociation, k_{off} ($s^{-1} \equiv Hz$), is much smaller than the frequency difference between the resonance signals ($\nu_{bound} - \nu_{free}$), the signals of the free and bound forms are observed separately. This range of k_{off} values corresponds to high strength of ligand – protein binding (see Table 1). The case $k_{off} \gg (\nu_{bound} - \nu_{free})$ corresponds to low strength of ligand binding; as a consequence, an averaged signal is observed. Between these limiting cases one deals with pronounced conformational broadening of signals with a coalescence point at $k_{off} \approx (\nu_{bound} - \nu_{free})$, which corresponds to moderately strong ligand binding (see Table 1).

Most experimental (including NMR) studies of protein – ligand interaction are carried out at different protein/ligand ratios. Often, NMR titration methods are used (see below). Consider how resonances of the free and bound forms of the ligand change depending on their ratio (*i.e.*, depending on the ligand/protein mole ratio). The NMR spectra calculated for strong, moderately strong and weak protein – ligand binding are shown in Fig. 5. As the ligand/protein mole ratio changes, either the intensities of the corresponding resonances (see Fig. 5 *a* for strong binding) or the resonance frequencies of the averaged signal (see Fig. 5 *b*,*c* for weak or moderately strong binding) change. In the case of moderately strong binding at ligand/protein mole ratios of the order

Table 4. Key parameters extracted from NMR spectra of protein – ligand complexes in solution.

Observed NMR parameters	Information extracted
Chemical shifts	The character of electron density distribution in the vicinity of atomic nuclei; the assignment of protein and ligand signals; secondary structure of proteins, proximity and orientation of aromatic fragments and(or) carbonyl groups; involvement of atoms in the formation of hydrogen bonds; mapping ligand binding sites
Integrated intensities of signals	The amount (proportion) of ligand and(or) protein, kinetic measurements
Spin – spin coupling constants	Characterization of chemical bonds (^1J and ^2J), torsion angles and relative populations of rotamers (^3J), position and strength of hydrogen bonds (^{2h}J and ^{3h}J) (see [a])
Nuclear Overhauser effects	Internuclear distances (including distances between magnetic nuclei of ligand and amino acid residues in the protein molecule), rotational diffusion rate of protein or ligand molecule, mobility of protein chain
Magnetization transfer rate	Rate constants for exchange between the free and bound forms of ligands, the size of the molecules
Relaxation times	The rate of rotational diffusion of protein or ligand molecules, the mobility of protein chain or ligands, parameters of conformational exchange
Residual constants of dipole – dipole interactions	Spatial orientation of chemical bonds, dynamic characteristics
Translational diffusion rate	The size and shape of molecules
The rate of exchange of mobile protons with deuterons	Arrangement of hydrogen bonds, the folding and unfolding rates of protein globule and(or) its fragments

[a] Parameters ^1J, ^2J and ^3J are the spin – spin coupling constants of atoms separated by one, two and three chemical bonds, respectively; parameters ^{2h}J and ^{3h}J are the spin – spin coupling constants of atoms separated by two and three chemical bonds, respectively, one of them being a hydrogen bond.

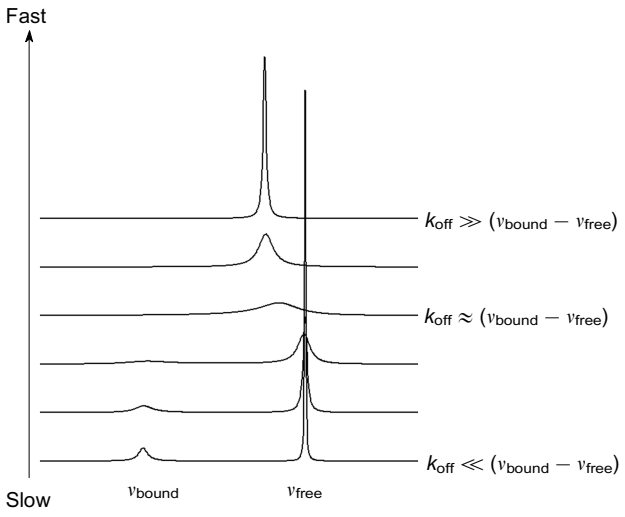


Figure 4. Simulated NMR spectra of a ligand at different strength of the protein–ligand complex. Calculated for the case of equilibrium between the free and protein-bound forms of the ligand.[28] The free/bound ligand ratio is 5 : 1. The natural linewidth (ignoring the contribution of chemical exchange) is 2 Hz for the free and 20 Hz for the protein-bound ligand. The frequency difference between the bound- (ν_{bound}) and free-ligand (ν_{free}) resonances is 300 Hz.

of unity the resonance lines can disappear owing to large broadeining.

Thus, the choice of particular NMR techniques for studying protein–ligand interactions depends strongly on the strength of ligand binding. Now we will dwell on the methods of investigation of strong and weak protein–ligand interactions.

3.1. NMR methods for studying tight protein–ligand complexes

The addition of a ligand to a protein resulting in the formation of a tight protein–ligand complex [$k_{off} \ll (\nu_{bound} - \nu_{free})$, see Fig. 5 *a*] leads to a duplication of signals in the NMR spectrum. One set of signals corresponds to the free form of the protein or ligand (depending on what is in excess) while the other set corresponds to the complex. At equimolar ligand/protein ratio, the NMR spectra exhibit only the signals of the protein–ligand complex. Structure determination of the complex can be performed using the full range of modern multidimensional heteronuclear NMR techniques (see Refs 79 and 80) and ^{13}C- and ^{15}N-enriched protein (or deuterated protein in the case of large biomolecules, see Ref. 81).

Carrying out structure determination of the protein–ligand complex requires identification of resonances of the bound ligand and the assignment of the nuclear Overhauser effects (NOE) between the ligand and protein signals. In particular, unambiguous assignment of the NOE at the protein–ligand interface assumes the use of ^{13}C- or ^{15}N-labelled ligand (Fig. 6 *a*). This allows one to measure the NOESY spectra with selection of protons coupled to ^{13}C or ^{15}N nuclei (in particular, 2D or 3D $^{13}C-^{1}H$ HSQC-NOESY spectra). For instance, the assignment of methyl groups of antibacterial drug Trimethoprim (TMP) in its complex with bacterial DHFR was reported (Fig. 7).[82, 83]

At the same time, labelling of small molecules with ^{13}C or ^{15}N isotopes is a time-consuming and expensive process. Often, it is more convenient and cheap to obtain proteins enriched with these stable isotopes using genetic engineering methods.[84] In this case, the NOE between the ligand containing mainly magnetically inactive ^{12}C nuclei and the ^{13}C-labelled protein (see Fig. 6 *b*) can be identified using isotope filtering.[85, 86] As a consequence, the

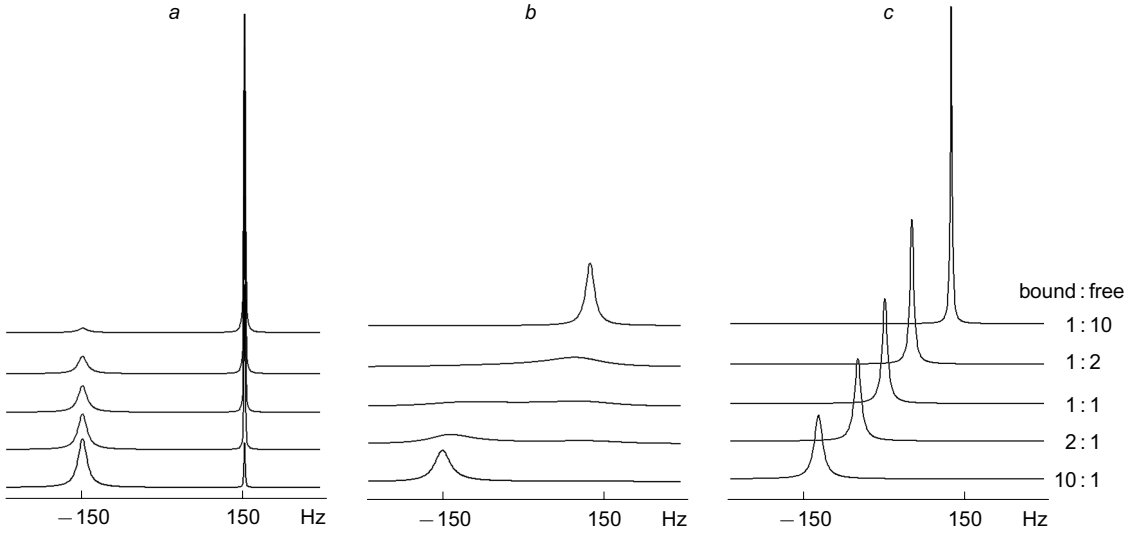


Figure 5. Ligand signals in simulated NMR spectra at different bound/free ligand ratios provided that both forms of the ligand are at equilibrium.[28]
Formation of tight protein–ligand complex, $k_{off} = 0.5\ s^{-1}$ (*a*); moderately strong protein–ligand binding, $k_{off} = 500\ s^{-1}$ (*b*); and low ligand affinity for protein, $k_{off} = 5 \times 10^5\ s^{-1}$ (*c*). The natural linewidth (ignoring conformational broadening effects) is 2 Hz for the free and 20 Hz for the protein-bound ligand.

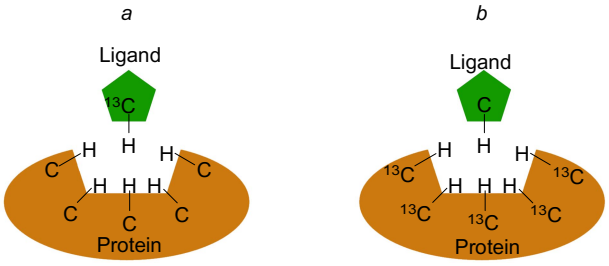


Figure 6. Types of isotope substitution for identification of nuclear Overhauser effects between the ligand and protein protons in the protein – ligand complex.[19]
When using isotopically enriched ligand (*a*), NMR experiments with isotope selection are carried out. In the case of totally isotopically labelled protein (*b*), an unlabelled ligand and isotope-filtered NMR experiments are used. In both cases, only $^{12}C-H \leftrightarrow H-^{13}C$ NOE are observed.

NOESY or TOCSY spectra will exhibit only signals of the protons bound to ^{12}C atoms, while signals of protons from $^{1}H-^{13}C$ pairs will be filtered out. If the isotope filtering is not applied on the indirect axis of the 2D spectrum, one can observe the NOE between the ligand and protein protons in the resulting NOESY spectrum. A fragment of the ^{13}C-filtered 2D NOESY NMR spectrum of the ternary complex of human ^{13}C,^{15}N-dihydrofolate reductase (DHFR) with coenzyme nicotinamide adenine dinucleotide phosphate (NADPH) and antibacterial drug Trimethoprim is shown in Fig. 8. An analysis of the cross-peaks in this spectrum made it possible to determine numerous NOEs between the protein and both ligands and to use them as a basis for calculating the 3D structure of the complex (see Fig. 7 *a*).[87]

The approaches mentioned above were employed for structure determination of a few hundred complexes of small ligands with proteins whose structures are available from the Protein Data Bank (http://www.rcsb.org). Three examples of the structures of the complexes of pharmacologically important proteins with small ligands are presented in Fig. 9.

NMR structure determination of protein – ligand complexes is simplified if the structures of complexes of this protein or its close analogues with other ligands are known. In this connection, a ligand docking method using restraints extracted from NMR data was proposed.[90] Yet another recently proposed technique, NMR^2 (Nuclear Magnetic Resonance Molecular Replacement), does not require the assignment of protein signals but needs information on the structure of the complex of a homologue.[91,92] This reduces the time

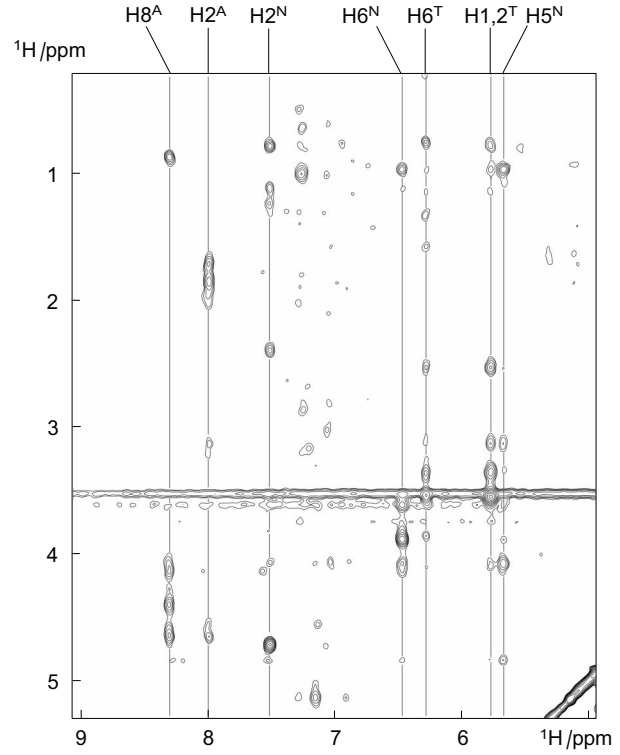


Figure 8. Fragment of ^{13}C-filtered 2D NOESY spectrum of a complex of human dihydrofolate reductase with coenzyme NADPH and antibacterial drug Trimethoprim.[87]
Vertical axis: proton signals of the ligands (NADPH and TMP) bound to nonmagnetic ^{12}C atoms. All $^{1}H-^{13}C$ signals of the labelled protein are suppressed. Positions of protons of the cofactor NADPH and inhibitor TMP are shown on the upper horizontal axis.

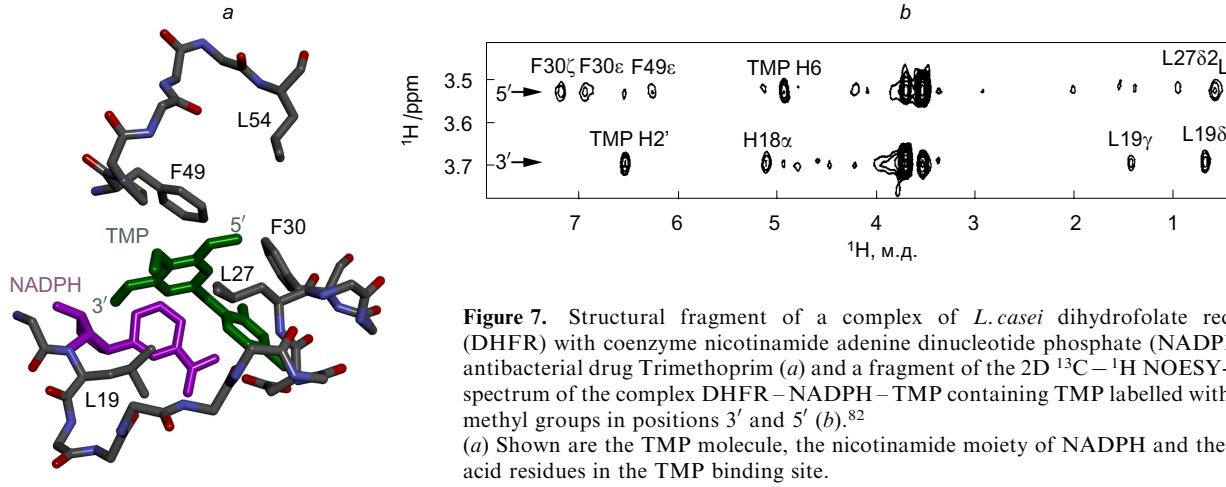


Figure 7. Structural fragment of a complex of *L. casei* dihydrofolate reductase (DHFR) with coenzyme nicotinamide adenine dinucleotide phosphate (NADPH) and antibacterial drug Trimethoprim (*a*) and a fragment of the 2D $^{13}C-^{1}H$ NOESY-HSQC spectrum of the complex DHFR – NADPH – TMP containing TMP labelled with ^{13}C at methyl groups in positions 3′ and 5′ (*b*).[82]
(*a*) Shown are the TMP molecule, the nicotinamide moiety of NADPH and the amino acid residues in the TMP binding site.

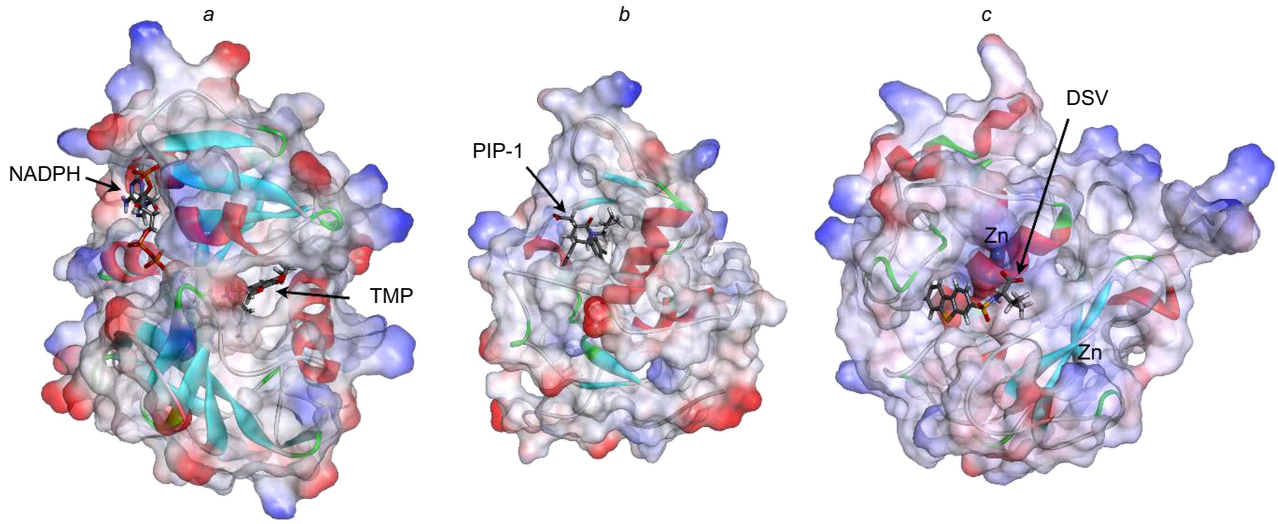


Figure 9. Solution NMR structures of tight protein–ligand complexes.
Ternary complex of human DHFR (186 amino acid residues, ~20 kDa) with coenzyme NADPH and antibacterial drug Trimethoprim [87] (*a*); binary complex of N-terminus of MDM-2 oncoprotein (125 amino acid residues, 14.1 kDa) with PIP-1 inhibitor of protein–protein (MDM-2 : p53) interaction [88] (*b*); and complex of the catalytic domain of human matrix metalloprotease MMP-12 (165 amino acid residues, 18.3 kDa) with enzyme inhibitor DSV [N-(Dibenzo[b,d]thiophene-3-sulfonyl)-valine] and two zinc ions [89] (*c*).

necessary to determine the structure of the new complex from a few months to a few days.

Determination of ligand binding constants in tight ligand–protein complexes using NMR spectra is a challenge since in this case one deals with quantitative ligand binding to the protein (see Table 1). However, it is possible to determine the ratio of the K_a constants for two competitively binding ligands and then obtain the unknown quantity provided that one K_a constant is known. If two ligands, L_1 and L_2, competitively bind to protein P, the binding constant ratio is given by

$$\frac{K_a^{L_1}}{K_a^{L_2}} = \frac{[PL_1]}{[PL_2]}\frac{[L_2]}{[L_1]} \tag{7}$$

Assuming initial equal molar concentrations of the protein and ligands

$$[P]_0 = [L_1]_0 = [L_2]_0, \tag{7a}$$

after equilibration one gets

$$\frac{[PL_1]}{[PL_2]} = \frac{[L_2]}{[L_1]} \tag{7b}$$

This equality holds for zero concentration of the free protein, being a good admission in the case of strongly binding ligands. Then Eqn (7) can be simplified:[93]

$$\frac{K_a^{L_1}}{K_a^{L_2}} = \left(\frac{[PL_1]}{[PL_2]}\right)^2 = \left(\frac{[L_2]}{[L_1]}\right)^2 \tag{8}$$

The expression for the binding constant ratio at arbitrary concentrations of the protein and two competitively binding ligands L_1 and L_2 can be derived:

$$\frac{K_a^{L_1}}{K_a^{L_2}} = \frac{[PL_1]}{[PL_2]}\left(\frac{[PL_1]}{[PL_2]}m + m - 1\right)\left(\frac{[PL_1]}{[PL_2]}(n-1) + n\right)^{-1} \tag{9}$$

where n and m respectively denote the ratios of the concentrations of the ligands L_1 and L_2 to the total concentration of the protein

$$[L_1]_0 = n[P]_0 \tag{10}$$

$$[L_2]_0 = m[P]_0 \tag{10a}$$

The concentration ratio of the complexes PL_1 and PL_2 or free ligands L_1 and L_2 can be estimated from the integrated intensity ratio of the corresponding characteristic signals in the NMR spectra. The binding constant of antitumour drug trimetrexate to DHFR determined by this method was $\sim 2 \times 10^8$ M^{-1}.[93]

3.2. NMR methods to study weak ligand binding to proteins

The most interesting application of NMR methods for the detection of weak protein–ligand interaction is to search for structural fragments that specifically interact with the target protein since small fragments usually have low binding constants. In the case of moderately strong or weak ligand binding to a protein [$k_{off} > (\nu_{bound} - \nu_{free})$, see Fig. 5 *b*,*c*], the NMR spectra exhibit signals with averaged parameters (chemical shifts, relaxation rates, *etc.*). For instance, the observed exchange-averaged chemical shift (δ_{obs}) is given by

$$\delta_{obs} = P_{PL}\delta_{PL} + (1 - P_{PL})\delta_{free} \tag{11}$$

where P_{PL} is the proportion of the bound form of the protein or ligand and δ_{PL} and δ_{free} are the chemical shifts of the signal from the protein–ligand complex and from the free form, respectively. This equation allows one to derive the expression for the proportion of the bound form P_{PL}, which can be calculated from the chemical shift δ_{obs} provided that the chemical shifts for the free form and for the complex are known, namely

$$P_{PL} = \frac{\delta_{obs} - \delta_{free}}{\delta_{PL} - \delta_{free}} = \frac{\Delta\delta_{obs}}{\Delta\delta_{max}} \quad (12)$$

where $\Delta\delta_{max}$ is the chemical shift difference between the complex and the free form.

Equations (11) and (12) are applicable to both the ligand signals at $\delta_{free} \equiv \delta_L$ and the protein signals at $\delta_{free} \equiv \delta_P$. Note that $P_{PL} = [PL]:[P]_0$ if the NMR experiments involve measurements of protein signals and $P_{PL} = [PL]:[L]_0$ if the ligand signals are measured.

3.2.1. Determination of ligand binding constants in the case of weak protein–ligand complexes

Starting with self-explanatory relationships $[P]_0 = [P] + [PL]$ and $[L]_0 = [L] + [PL]$, expression (2) for the dissociation constant K_d can be transformed with ease into a quadratic equation with the following solution for [PL]:

$$[PL] = \frac{1}{2}\left\{(K_d + [P]_0 + [L]_0) - \left[(K_d + [P]_0 + [L]_0)^2 - 4[P]_0[L]_0\right]^{1/2}\right\} \quad (13)$$

Equations (13) and (12) should be used together in order to evaluate the proportion of the bound form of the protein from the chemical shifts measured in the NMR titration experiments. Then, the K_d values can be determined using nonlinear regression methods. For instance, if protein signals are measured under titration with the ligand, one has

$$\Delta\delta_{obs} = \frac{\Delta\delta_{max}}{2[P]_0}\left\{(K_d + [P]_0 + [L]_0) - \left[(K_d + [P]_0 + [L]_0)^2 - 4[P]_0[L]_0\right]^{1/2}\right\} \quad (14)$$

An analysis of changes in the chemical shifts in the NMR titration experiments gives not only K_d, but also the parameter $\Delta\delta_{max}$. The changes in the chemical shifts of the protein ^{1}H signal during titration with ligands having binding constants in the range from 10^3 to 10^7 M^{-1} calculated using formula (14) are demonstrated in Fig. 10.

As can be seen, different patterns of the changes in the chemical shifts are only observed for the binding constants of at most 10^6 (*i.e.*, for ligands with $K_d > 1$ μM). Most often, the protein concentration is maintained at a certain constant level during NMR titration experiments, the ligand concentration being varied. Here, one should take account of the decrease in the total protein concentration with increasing volume of the solution during the titration process. An alternative approach involves variation of both ligand and protein concentrations.[94] This makes it possible to increase the accuracy of the ligand dissociation constant.

Most often, the chemical shifts of particular atoms in the protein molecule are measured in the NMR titration experiments using 2D heteronuclear spectroscopy. In this case, one can detect the chemical shifts of protons and heavier nuclei (^{15}N or ^{13}C) for each signal whose position varies upon adding a ligand. The constant K_d can be calculated using parameters corresponding to the weighted chemical shifts $\delta(^1H)$ and $\delta(^{13}C)$ or $\delta(^1H)$ and $\delta(^{15}N)$ for each protein signal under study.[73]

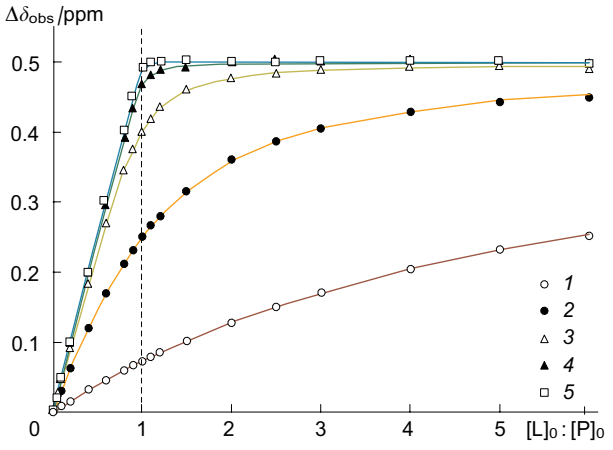


Figure 10. Changes in ^{1}H chemical shifts of a protein during titration with increasing amount of ligand. The chemical shifts were calculated for 1 : 1 binding stoichiometry, initial protein concentration of 0.2 mM and titration with the ligand solution (ligand concentration is 4 mM) assuming $\Delta\delta_{max} = 0.5$ ppm for the protein ^{1}H signal under study. Calculations were carried out with inclusion of the increase in the solution volume in the course of titration process, which leads to a decrease in the total concentrations of the protein and ligand. Calculated using Eqn (14) for ligand binding constants of 10^3 (*1*), 10^4 (*2*), 10^5 (*3*), 10^6 (*4*) and 10^7 M^{-1} (*5*).

Not only nonlinear regression methods, but also graphical methods based on conversion of experimental data to more convenient systems of coordinates can be used to analyze the results of NMR titration experiments. For instance, the Scatchard transformation has the form[95]

$$K_a = \frac{1}{K_d} = \frac{[PL]}{[L][P]} = \frac{[PL]}{[L]([P]_0 - [PL])} \quad (15)$$

$$\frac{[PL]}{[L]} = -\frac{[PL]}{K_d} + \frac{[P]_0}{K_d} \quad (16)$$

If the ligand is present in excess, one has $[L] \approx [L]_0$. Using expression (11), one can go from Eqn (16) to the experimentally measured parameters of the NMR spectrum:

$$\frac{\Delta\delta_{obs}}{[L]_0} = -\frac{\Delta\delta_{obs}}{K_d} + \frac{\Delta\delta_{max}}{K_d} \quad (17)$$

As an example, Fig. 11 shows the Scatchard plot of changes in the characteristic signals of histidine Hβ residue during titration of a fragment of human β-amyloid peptide with $ZnCl_2$ solution in the course of formation of a complex between the peptide and zinc ion.[96] Nonlinear portion of the plot is observed at high Zn^{2+} concentrations, which is due to the deviation of the binding stoichiometry from 1 : 1.

The results of NMR titration experiments are often analyzed using the Benesi–Hildebrand and Scott coordinates.[97] In the former case, the plots of $(\Delta\delta_{obs})^{-1}$ *vs.*

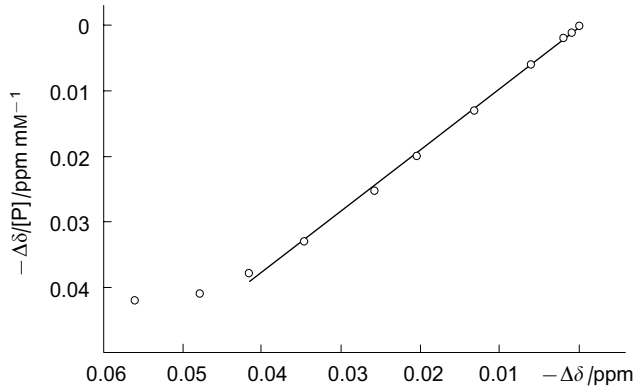


Figure 11. Changes in the chemical shift of Hβ signal of the His6 amino acid residue of isoD7-Aβ_{1-10} fragment of human β-amyloid peptide during titration with $ZnCl_2$ solution plotted in the Scatchard coordinates.[96]
The initial peptide concentration is 1.0 mM and the maximum $[Zn^{2+}]$:[peptide] mole ratio is 10. The calculated dissociation constant of zinc ion from the complex with the peptide is 0.84 ± 0.04 mM.

$([L]_0)^{-1}$ are constructed; in the latter case, the $[L]_0 : \Delta\delta_{obs}$ ratio is plotted *vs.* $[L]_0$.

$$\frac{1}{\Delta\delta_{obs}} = \frac{K_d}{[L]_0 \Delta\delta_{max}} + \frac{1}{\Delta\delta_{max}} \quad (18)$$

$$\frac{[L]_0}{\Delta\delta_{obs}} = \frac{K_d}{\Delta\delta_{max}} + \frac{[L]_0}{\Delta\delta_{max}} \quad (19)$$

However, the Scatchard, Benesi–Hildebrand and Scott transformations are no longer routinely used since nonlinear regression methods for direct analysis of Eqn (14) are now available. Software packages and algorithms for 1D and 2D total lineshape analysis of NMR signals under exchange between the free and bound forms of protein and ligand were developed (for 1D spectra, see Refs 94 and 98; for 2D spectra, see Refs 99–101).

3.2.2. Methods for the determination of the rate constant for ligand dissociation from it complex with the target protein

NMR spectroscopy techniques can be used to determine the rate constant for ligand dissociation k_{off}. As mentioned above, this parameter along with K_d characterizes the strength of the ligand–target binding and specifies the residence time of the ligand in the bound state, $\tau = 1/k_{off}$.[102] Measurements of k_{off} are possible since the width of the ligand signal depends on this quantity (see Figs 4 and 5). The rate constants for ligand dissociation were determined using traditional methods of total lineshape analysis for 1D or 2D NMR spectra.[100, 103, 104] Quite recently, it was proposed to determine the k_{off} values in the NMR relaxation dispersion experiments. We will briefly outline the principles of this method.

The width of the NMR signal is governed by the transverse relaxation rate $R_2 = 1/T_2$. Most often, this parameter is measured using the Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence representing a train of equidistant 180-degree pulses (Fig. 12).

Each 180-degree pulse creates a spin echo signal with the retention of coherence in the plane perpendicular to the direction of external magnetic field. By measuring the dependence of the intensity of the NMR signal on variable number of the 180-degree pulses and(or) on the time delay between them one can calculate the transverse relaxation rate. However, if chemical exchange occurs between states with different resonance frequencies ν_A and ν_B, refocussing of the spin echo signals becomes impossible since the precession rate changes upon transition of the atomic nucleus from the state A to the state B and *vice versa*. Note that the longer the time delay between the 180-degree pulses the higher the probability of these transitions. Therefore, in the system with chemical exchange the R_2^{obs} values measured at different time delays between the 180-degree pulses of the CPMG sequence will be different. Usually, relaxation dispersion measurements are carried out with fixed duration T of the CPMG pulse sequence and variable number of the 180-degree pulses (see Fig. 12). For each CPMG sequence one can obtain the effective radiofrequency (rf) field strength (ν_{CPMG}) determined by the duration of the time delays (τ) between the 180-degree pulses[106]

$$\nu_{CPMG} = \frac{1}{2\tau} \quad (20)$$

The apparent (effective) transverse relaxation rate R_2^{obs} for each ν_{CPMG} value can be estimated from the ratio of the signal intensity measured after application of the CPMG pulse sequence to the signal intensity measured immediately after application of the first 90-degree pulse,

$$I_T^{\nu_{CPMG}} = I_0 e^{-TR_2^{obs}} \Rightarrow R_2^{obs} = -T^{-1} \ln\left(\frac{I_T^{\nu_{CPMG}}}{I_0}\right) \quad (21)$$

The parameter R_2^{obs} depends on the rate constant for chemical exchange k_a, on the resonance frequency differ-

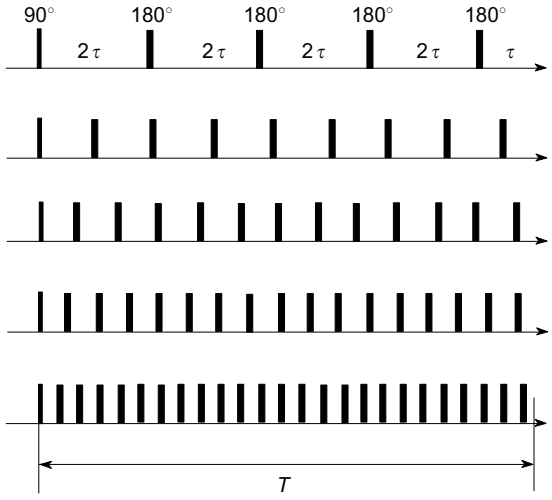


Figure 12. Pulse scheme employed to measure relaxation dispersion profiles using the GPMG pulse sequence.[105]
A series of experiments involves measurements with increasing number of refocussing 180-degree pulses within constant-time interval T of the GPMG pulse sequence. The effective rf field strength increases with increasing number of pulses.

ence, $\Delta\omega$, between the states A and B and on the time delay τ_{CPMG}

$$R_2^{obs} = R_2 + k_a + k_a \frac{\sin(\Delta\omega, \tau_{CPMG})}{\Delta\omega\, \tau_{CPMG}} \quad (22)$$

where R_2 is the transverse relaxation rate in the absence of chemical exchange.

Figure 13 illustrates chemical exchange between the states A and B occurring at different rates and presents the plots of R_2^{obs} values measured in the model experiments (see above) *vs.* CPMG field strength. The curve calculated using $k_a = 0$ corresponds to the absence of chemical exchange. In this case the linewidth is determined by the parameter R_2 only. At nonzero exchange rates, the linewidth increases by k_a.[106] As can be seen, the linewidths vary only slightly with increasing k_a; however, the R_2^{obs} values appear to be much more informative. It should be noted that the resonance frequency difference $\Delta\omega$ can be determined by measuring the intensity of a single line, A or B. However, this is a difficult task involving a series of measurements at low field strength ν_{CPMG} (*i.e.*, at long time delays τ_{CPMG}); in most cases this is impossible because at least one 180-degree pulse should be applied within the time T.

The principles underlying the NMR relaxation dispersion measurements can be used to study the ligand exchange between the free and protein-bound forms.[107] In this case the rate constant for exchange, k_{ex}, which can be determined in the relaxation dispersion experiments, is related to the constants k_{on} and k_{off} for equilibrium (1) between the free and protein-bound forms of the ligand:

$$k_{ex} = k_{off} + k_{on}[P] = k_{off}\left([P]\, K_d^{-1} + 1\right) \quad (23)$$

The binding of ^{15}N-labelled peptide to its target, human prothrombin, was studied by measuring the relaxation dispersion of ^{15}N nuclei in amide groups using heteronuclear NMR spectroscopy methods.[108] The authors not only measured the rate of peptide dissociation from its complex with prothrombin, but also revealed the existence of two conformations of the bound peptide, which undergo interconversion at a rate of $\sim$100 s^{-1}. The kinetic parameters of ligand – target binding were determined from the analysis of ^{1}H (^{15}N, ^{13}C, ^{19}F) NMR relaxation dispersion experiments with the magnetic nuclei in ligands.[109 – 112]

It was proposed[113] to carry out relaxation dispersion measurements in the rotating frame $R_{1\rho}^{obs}$, *i.e.*, to use spin-locking rather than the CPMG sequence to hold the magnetization in the perpendicular plane. The authors emphasized that the spin-locking allows one to detect higher rates of ligand dissociation. Like the CPMG-experiments, this approach also requires a series of measurements at different spin-locking rf field strength (ω_{eff}) values. To measure the $R_{1\rho}^{obs}$ value, two NMR spectra are recorded, one at very short ($\sim$5 ms) and the other at long (from 600 to 1000 ms) spin lock time. The parameter $R_{1\rho}^{obs}$ depends on ω_{eff}:

$$R_{1\rho}^{obs} = R_2 + \left[P_L P_{PL} (\delta_L - \delta_{PL})^2 \omega_0^2\right] \left(k_{off} + \frac{\omega_{eff}^2}{k_{off}}\right)^{-1} \quad (24)$$

Here P_L and P_{PL} are respectively the proportions of the free and bound forms of the ligand, δ_L and δ_{PL} are the chemical shifts of the free-ligand and bound-ligand resonances, respectively, and ω_0 is the Larmor frequency. Having done a series of $R_{1\rho}^{obs}$ measurements at different ω_{eff} values, one can calculate the rate constant for ligand dissociation k_{off}.

3.2.3. Determination of protein – ligand binding stoichiometry

Prior to determining the binding constant, one should get the information on the protein – ligand binding stoichiometry. NMR spectroscopy along with isothermal titration calorimetry and spectrophotometric methods can be efficiently used for determination of stoichiometry. Usually, one uses the method of isomolar titration, or Job's method, which was initially proposed to determine the complexation stoichiometry of inorganic compounds by

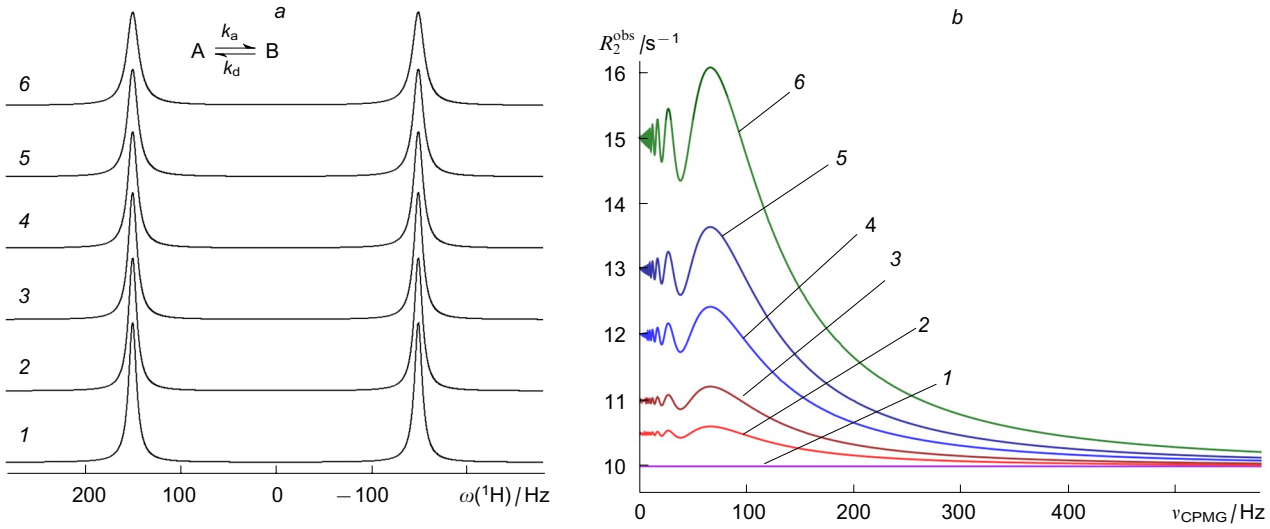


Figure 13. Chemical exchange between states A and B.[106]
(*a*) Spectra of the A and B states calculated for different exchange rates; $\Delta\omega = 300$ Hz, $T_2 = 0.1$ s; (*b*) relaxation dispersion (R_2^{eff}) plots *vs.* ν_{CPMG} (CPMG field strength); $\Delta\omega = 300$ Hz, $R_2 = 10$ s^{-1}, calculated using Eqn (22) for $k_a = 0$ (*1*), 0.5 (*2*), 1 (*3*), 2 (*4*), 3 (*5*) and 5 s^{-1} (*6*).

spectrophotometric methods.[114] The technique is based on the measurements of NMR spectral parameters (chemical shifts, signal intensities or other characteristic parameters listed in Table 4) for a series of samples with variable proportion of the ligand (υ) provided that the total concentration of the protein and ligand in each sample is constant

$$\upsilon = \frac{[L]_0}{[L]_0 + [P]_0} \tag{25}$$

$$[L]_0 + [P]_0 = C \tag{25a}$$

Most often, this is achieved by mixing variable volumes of stock solutions of the protein and ligand of the same concentration when preparing constant-volume samples. Clearly, the proportion of the ligand can vary from 0 to 1. Generally, the protein–ligand interaction can be described by the equation

$$m\,P + n\,L \rightleftharpoons P_mL_n \tag{26}$$

The concentration of the complex P_mL_n being formed can be expressed through υ and C (see Ref. 115):

$$\frac{[P_mL_n]}{K_a} = [P]^m[L]^n = \tag{27}$$

$$= \{C(1-\upsilon) - m[P_mL_n]\}^m \left(C\upsilon - n[P_mL_n]\right)^n$$

The maximum proportion of the complex P_mL_n at different ligand content is achieved at $d([P_mL_n])/d\upsilon = 0$. Differentiation of Eqn. (27) with respect to υ gives the expression for the mole fraction of the ligand which corresponds to the maximum concentration of the complex,[116]

$$\upsilon_{max} = \frac{n}{m+n} \tag{28}$$

Since the concentration of the complex is proportional to the product of $\Delta\delta_{obs}$ by the total concentration of the protein, $[P]_0$ [see Eqn (11) and comments], the binding stoichiometry can be determined using the plot of $\Delta\delta_{obs}[P]_0$ *vs.* mole fraction of the ligand. Information on the amount of the complex formed can be extracted from not only chemical shifts, but also signal intensities; this is useful if tight protein–ligand complexes are formed.

The calculated isomolar titration curves for a number of protein–ligand complexes of different stoichiometry are presented in Fig. 14 *a*.

It should be noted that expression (28) does not include binding constant, *i.e.*, the maximum value is attained at any binding constant value that allows one to identify the complexation-induced changes in the chemical shifts, signal intensities or other NMR spectral parameters. However, the shape of the curve depends on the binding constant. For instance, the isomolar titration curves calculated for five K_a values lying in the range from 10^3 to 10^7 M^{-1} (see Fig. 14 *b*) pass through a maximum at $\upsilon = 0.5$. The maximum is clearly seen if a tight protein–ligand complex ($K_a > 10^6$ M^{-1}) is formed and becomes less pronounced for weak ligand binding.

The method of isomolar NMR titration can be used to determine the stoichiometry of the interaction of proteins, peptides and other biomolecules with various ligands including metal ions. As an example, Fig. 15 demonstrates the results obtained in studies of the interaction between zinc ions and the metal-binding fragment of human β-amyloid peptide containing the so-called English (H6R) mutation.[96, 118]

The addition of zinc ions to the peptide H6R-Aβ(1-16) results in a mixture of two peptide–zinc complexes, one with a relatively low and the other with a higher metal binding constant. The NMR spectra of the first complex exhibit signals with the chemical shifts averaged between the free and bound forms of the peptide, while the spectra of the second complex demon-

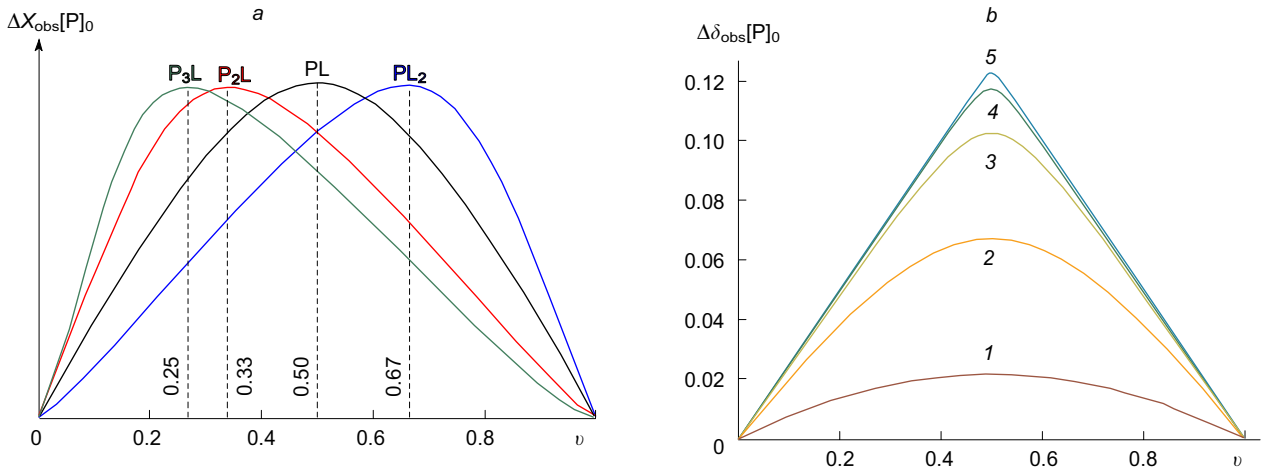


Figure 14. Isomolar titration curves calculated for protein–ligand complexes of different stoichiometry (*a*) and for a binary PL complex using different ligand binding constants (*b*).[117]
(*a*) ΔX_{obs} is the change in the NMR spectral parameter (chemical shift, signal intensity of the complex, relaxation time, *etc.*) on going from free to bound protein; (*b*) the $\Delta\delta_{obs}$ values calculated for $\Delta\delta_{max} = 0.5$ ppm and $[P]_0 + [L]_0 = 0.5$ mM. $K_a = 10^3$ (*1*), 10^4 (*2*), 10^5 (*3*), 10^6 (*4*) and 10^7 M^{-1} (*5*).

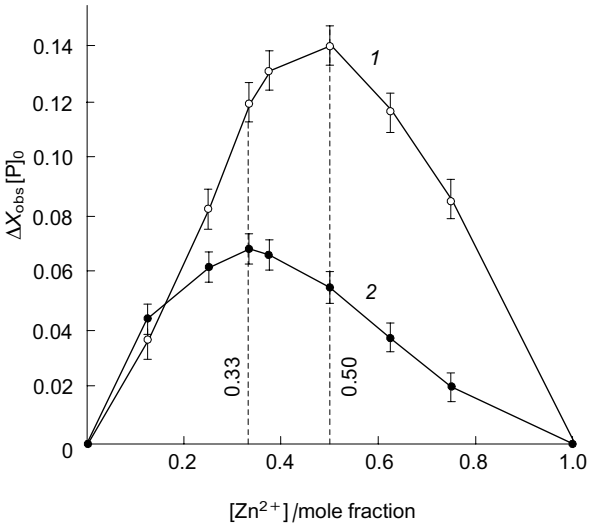


Figure 15. Results of the Zn^{2+} isomolar NMR titration of the metal-binding domain of β-amyloid peptide containing the so-called English pathogenic mutation (H6R).[118]
The interaction of the peptide with Zn^{2+} ions results in monomeric and dimeric forms of the complex; curves *1* and *2* correspond to the former and latter, respectively. The ordinate axis shows the product $\Delta X[P_0]$, where ΔX is either the change in the relative integrated intensity, ΔI, of the Val12 Hγ signal at 0.2 ppm corresponding to zinc-induced peptide dimer (curve *2*) or the change in the chemical shifts, $\Delta\delta$, of histidine Hδ2 signals observed upon the formation of monomeric peptide – zinc complex (curve *1*). The stoichiometry of the peptide – zinc interaction is 1 : 1 for curve *1* and 2 : 1 for curve *2*.

strate additional signals.[118] The concentration of the bound form of peptide can be determined from the changes in the chemical shifts for the first complex and from the integrated intensities of the new signals for the second complex. Isomolar NMR titration experiments revealed the formation of a peptide complex with one zinc ion in the former and a zinc-bound peptide dimer in the latter case.[96]

4. NMR screening methods

As mentioned above, NMR spectroscopy allows one to identify specific binding of small molecules to biomolecular targets of potential drugs even if the affinity of the interaction is low. This is particularly useful for the methods involving fragment-based screening. Identification of the interaction of low-molecular-mass compounds with the protein target is based on the detection of changes in the chemical shifts of the signals from the protein or from the ligands binding to it, in the relaxation and cross-relaxation rates of atomic nuclei, in the translational diffusion constant and other parameters that are sensitive to the formation of protein – ligand complexes.

4.1. Differences between NMR spectral parameters of large targets and small molecules

Many NMR spectral parameters of small molecules are significantly different from those of high-molecular-mass target proteins. The majority of modern NMR screening methods are based on these differences. A number of key parameters used in NMR studies of small and large molecules are listed in Table 5.

The NMR spectra of small molecules are characterized by narrow resonance lines and a relatively small range of changes in the ^{1}H chemical shifts for most functional groups. Contrary to this, the NMR spectra of proteins exhibit broad signals with half-widths ($\Delta\nu_{1/2}$) of tens (and even hundreds for large proteins) Hertz. Besides, the chemical shifts of signals of structured proteins change in a much wider range compared to the changes in the chemical shifts of corresponding groups in small molecules due to the effect of ring currents of closely spaced carbonyl groups and aromatic rings. For instance, ^{1}H resonances of proteins are often observed both in the region of negative proton chemical shifts and in the downfield region ($\delta > 10$). Differences in the linewidths of low- and high-molecular-mass compounds originate from large differences between corresponding transverse relaxation rates.

The longitudinal (T_1) and transverse (T_2) relaxation times of ^{1}H nuclei are plotted *vs.* correlation time of rotational diffusion of molecules (τ_c) in Fig. 16 *a*. On going from low-molecular-mass compounds ($10^{-12} < \tau_c < 10^{-10}$ s) to high-molecular-mass ones ($10^{-9} < \tau_c < 10^{-7}$ s) the transverse relaxation time T_2 decreases by four to five orders of magnitude. The NMR linewidths increase correspondingly. Even more pro-

Table 5. Main differences between NMR spectral parameters of low-molecular-mass compounds and high-molecular-mass biomolecules.

^{1}H chemical shifts	Transverse relaxation rate	^{1}H – ^{1}H nuclear Overhauser effect	Magnetization transfer rate	Rate constant for translational diffusion
		Low-molecular-mass compounds		
Typically, the chemical shift values span the range from ~ +1 to ~ +9 ppm for most compounds	Low (narrow signals; linewidth less than 1 Hz)	Positive values (the limiting value is +0.5)	Low rate constants for cross-relaxation and spin diffusion	High
		High-molecular-mass protein targets		
Wider range of chemical shift values, from −2 to +12 ppm for proteins and to +15 ppm for nucleic acids	High (linewidths of the order of tens of Hz)	Negative values (the limiting value is −1.0)	High rate constants for cross-relaxation and spin diffusion	Low

nounced differences are observed for the cross-relaxation rates and the Overhauser effects related to them. Figure 16,*b* presents the plots of the cross-relaxation rates (σ) between ^{1}H nuclei and the NOE *vs.* τ_c. Assuming isotropic rotational motion of a molecule with the correlation time τ_c, the cross-relaxation rate (σ_{ij}) between protons i and j depends on the internuclear distance (r_{ij}) and on the Larmor frequency (ω) as follows:

$$\sigma_{ij} = \frac{\gamma^4 (h/2\pi)^2}{10 r_{ij}^6} \left(\frac{6\tau_c}{1 + 4\omega^2 \tau_c^2} - \tau_c \right) \qquad (29)$$

where γ is the gyromagnetic ratio of proton.

Small molecules ($\tau_c < 0.2-0.4$ ns) are characterized by positive NOEs and σ_{ij} (see inset in Fig. 16 *b*).

As the correlation time τ_c increases, the NOE quite rapidly reaches a limiting value of -1.0 (at $\tau_c \approx 2$ ns). At long correlation times τ_c (*i.e.*, $\omega_0\tau_c \gg 1$), the first term in the rightmost bracketed expression in Eqn (29) becomes negligible and the parameter σ_{ij} is thus proportional to $-\tau_c$; therefore, the absolute values of the cross-relaxation rates continuously increase with increasing τ_c (*i.e.*, with increasing molecular mass of the protein). This leads to an increase in the rate of magnetization transfer between ^{1}H nuclei by the cross-relaxation mechanism, *i.e.*, to more efficient spin diffusion. Thus, spin diffusion appears to be highly efficient mechanism of magnetization transfer between nuclei in large biomolecules containing numerous closely spaced pairs of protons. In the case of small molecules this phenomenon can be ignored.

The relaxation properties of molecules described above depend on the rates of their rotational motions. At the same time, going from low-molecular-mass to high-molecular-mass compounds also causes changes in the translational diffusion rate; this parameter can be measured in the DOSY experiments.[120–122]

Upon the binding to large target, a small ligand acquires all properties of the large, *i.e.*, its chemical shifts change, the resonance lines are broadened, the cross-relaxation rate increases considerably, while the rate of translational diffusion decreases. This allows one to distinguish between signals from the binding and non-binding ligands. Binding of a ligand to a biomolecule can be detected by monitoring changes in the NMR signals of either the target or ligand. These two groups of approaches serve a basis of modern NMR screening methods.

4.2. NMR screening methods based on detection of protein signals

Chemical shifts represent the main source of information on the ligand binding when the protein signals are detected. Complexation of biomolecules with small ligands is accompanied by insignificant changes in the parameters determined by the rates of rotational or translational motions. However, the chemical shifts provide valuable information when using biomolecules labelled with the stable isotopes ^{13}C and(or) ^{15}N.

4.2.1. Mapping the ligand binding site using changes in chemical shifts

Ligand-induced changes in the chemical shifts of biomolecules are usually monitored using 2D heteronuclear correlation NMR spectra (HSQC or TROSY). The addition of unlabelled ligand to a protein enriched by the stable isotopes ^{13}C and(or) ^{15}N can cause the changes of chemical shifts of protein signals. This can be illustrated using a series of 2D ^{15}N–^{1}H HSQC spectra of a complex of *L. casei* dihydrofolate reductase with antibacterial drug Trimethoprim measured during titration with inorganic phosphate (Fig. 17).

Most signals retain their position as the concentration of phosphate ions increases; however, certain resonance from protein backbone or side-chain amide groups (the assignment is indicated in Fig. 17) exhibit considerable changes in the ^{1}H and(or) ^{15}N chemical shifts.

Dihydrofolate reductase (DHFR) has two binding sites, one for the substrate (or inhibitor, as in this case)

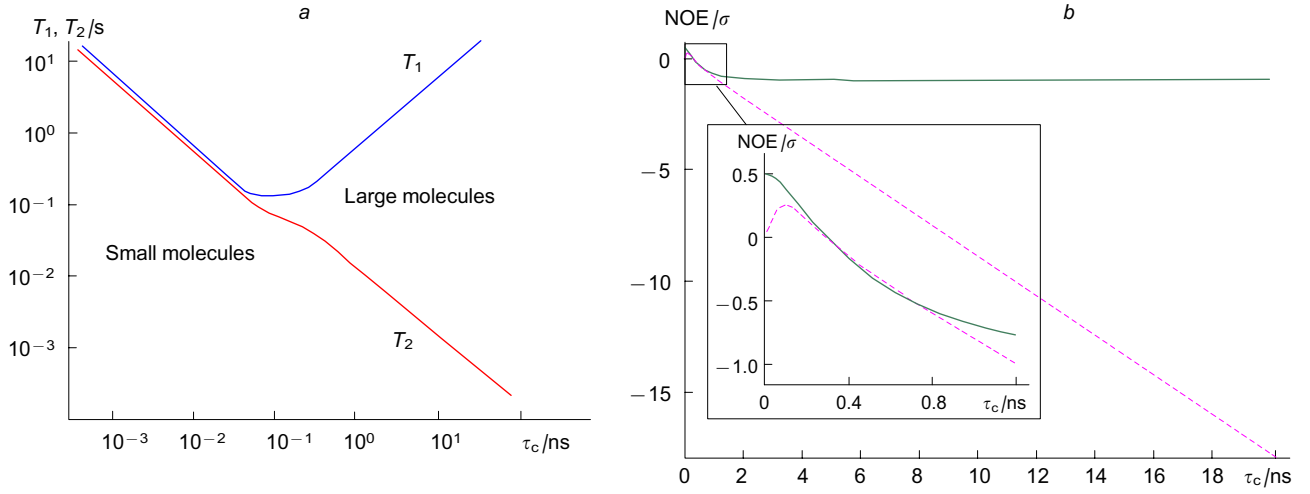


Figure 16. Longitudinal and transverse relaxation times of ^{1}H nuclei (*a*) and the NOE (solid line) and cross-relaxation rate (dashed lines) between two ^{1}H nuclei separated by a distance of 2 Å plotted *vs.* correlation time of rotational diffusion of the molecule (*b*). Calculated using Eqn (29) and expressions for T_1 and T_2 (see Refs 22 and 119) assuming isotropic rotation of the molecule and a proton resonance frequency of 600 MHz.

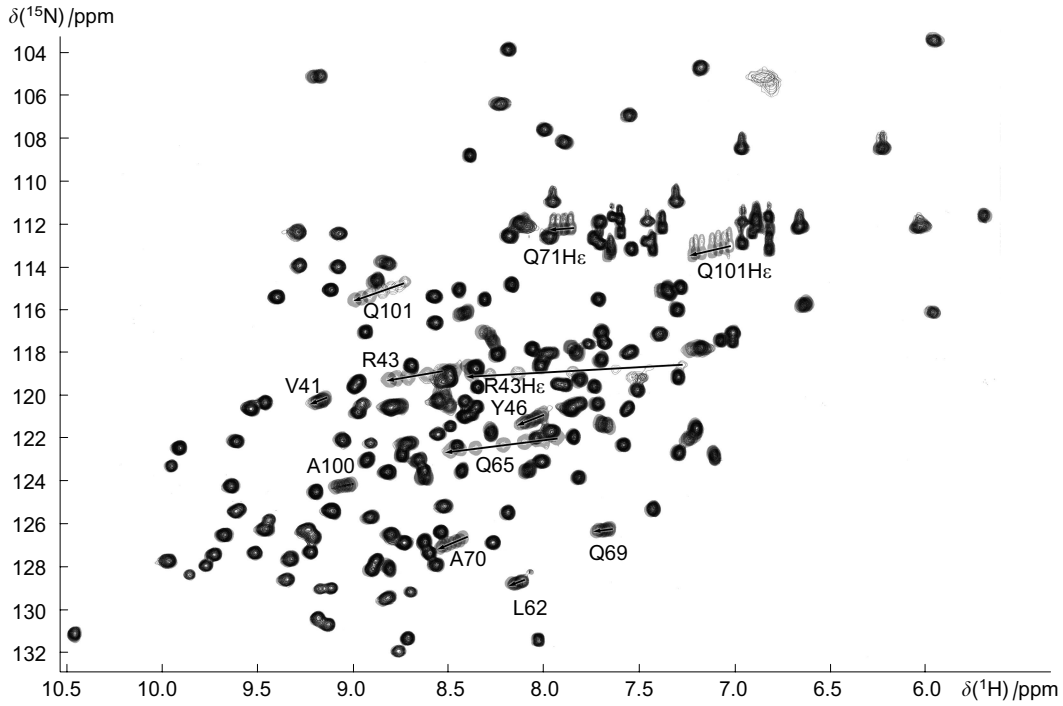


Figure 17. A series of 2D $^{15}N-^{1}H$ HSQC spectra of a binary complex of *L. casei* DHFR with antibacterial drug Trimetoprim measured during titration with inorganic phosphate (the phosphate : protein mole ratio increased from 0 to 20). The protein concentration was 2 mM. Shown are the signals of protein backbone and side chain amide groups, which change the chemical shift upon the interaction with phosphate ions.

and the other for the coenzyme NADPH containing three phosphate groups.[83] The binary complex of the enzyme with TMP used in the NMR titration experiments had unoccupied coenzyme binding site. Having specified the amino acid residues that noticeably changed their positions with increasing concentration of the inorganic phosphate, one can see that almost all of them surround three phosphate groups of the missing ligand, NADPH (Fig. 18).

This example clearly illustrates the potential of NMR spectroscopy in the detection of binding between small fragments of larger ligands and proteins. In spite of low binding constant of inorganic phosphate to DHFR, NMR titration experiments enable quite accurate mapping of the environment of phosphate groups within NADPH molecule (interaction of protein with this coenzyme was optimized evolutionarily).

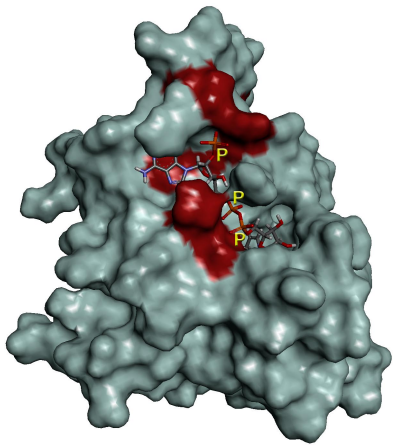


Figure 18. Structure of ternary complex of *L. casei* DHFR with antibacterial drug Trimetoprim and coenzyme NADPH.[83]
Shown is the NADPH molecule in the binding site of the protein (shown as the molecular surface). Positions of three phosphate groups of NADPH are specified. The DHFR amino acid residues characterized by considerable changes in the ^{1}H and ^{15}N chemical shifts of amide groups in the $^{15}N-^{1}H$ HSQC spectra of the binary complex of *L. casei* DHFR with TMP during titration with inorganic phosphate are given in red.

4.2.2. The SAR by NMR method

Mapping the binding site of the ligand to the target protein based on monitoring of changes in the chemical shifts (see above) underlies establishment of structure–activity relationships by NMR spectroscopy (SAR by NMR). The principle of this method (Fig. 19) was formulated in 1996 by the researchers from the Stephan Fesik group.[123]

The binding site of a small molecule is mapped by chemical shift changes for particular amino acid residues of the protein in the NMR titration experiments. Measurements of $^{15}N-^{1}H$ or $^{13}C-^{1}H$ HSQC spectra usually take at most 10–30 min. If a protein has two pockets (*1* and *2* in Fig. 19 *a*,*b*) separated by a relatively short distance, two different fragments binding in these pockets can be identified after two series of titration experiments (see Fig. 19 *e*,*f*). Fragment libraries for NMR screening experiments are usually created using the results of virtual screening. Successful implementation of the SAR

by NMR method requires information on the assignment of protein signals in the heteronuclear correlation spectra and on the 3D structure of the protein. Having identified both fragments, one can optimize their structures to increase the efficiency of binding (see Fig. 19 *c*). Eventually, the two fragments should be linked appropriately (see Fig. 19 *d*). The molecule thus designed contains two affinity fragments, which creates prerequisites for high binding constant of the new compound to the protein.

Numerous examples of successful application of the SAR by NMR method are available. Some of them will be demonstrated in Section 6 of this review. Here we present a schematic of the discovery of an inhibitor of the matrix zinc-dependent metalloprotease stromelysin-3 (Fig. 20)[124] as one of the first examples. The lead design took less than six months. First, two small fragments were identified using NMR titration experiments (see Fig. 20). One fragment binds to the zinc atom in the active site of enzyme and has a relatively low affinity (K_d = 17 mM). The second, larger, fragment binds in the adjacent pocket and has a dissociation constant of 20 μM. Structure optimization of the system comprising two fragments and a linker led to a compound with a dissociation constant of 15 nM and much higher bioavailability compared to the previously found peptide inhibitors of stromelysin-3.[124]

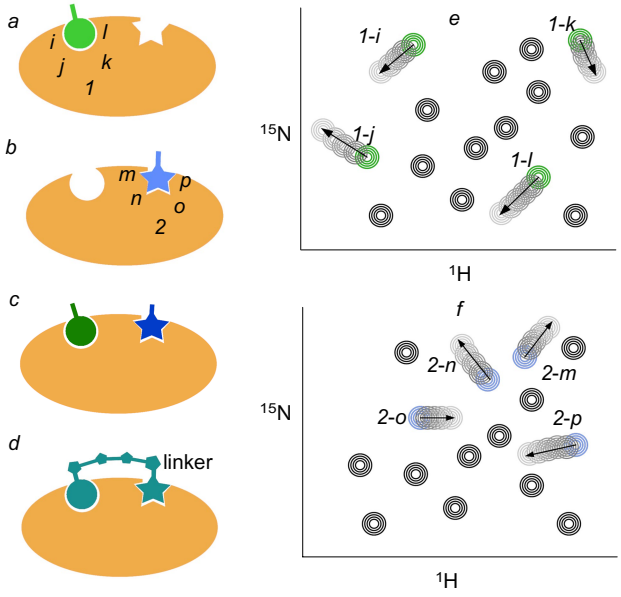


Figure 19. The SAR by NMR method: an outline.[123]
Identification of the first (*a*) and second (*b*) fragments binding to the target pockets *1* and *2*, respectively, using NMR titration methods (spectra *e* and *f*); structure optimization of the identified fragments aimed at increasing their affinity for the target protein (*c*) and bridging two fragments by the linker providing high binding constant of the newly designed molecule (*d*).
Spectra *e* and *f* shown the resonance lines of the amino acid residues *i*–*l* surrounding pocket *1* (given in green) and *m*–*p* surrounding pocket *2* (given in blue). These lines change their positions when the concentration of the corresponding fragment increases.

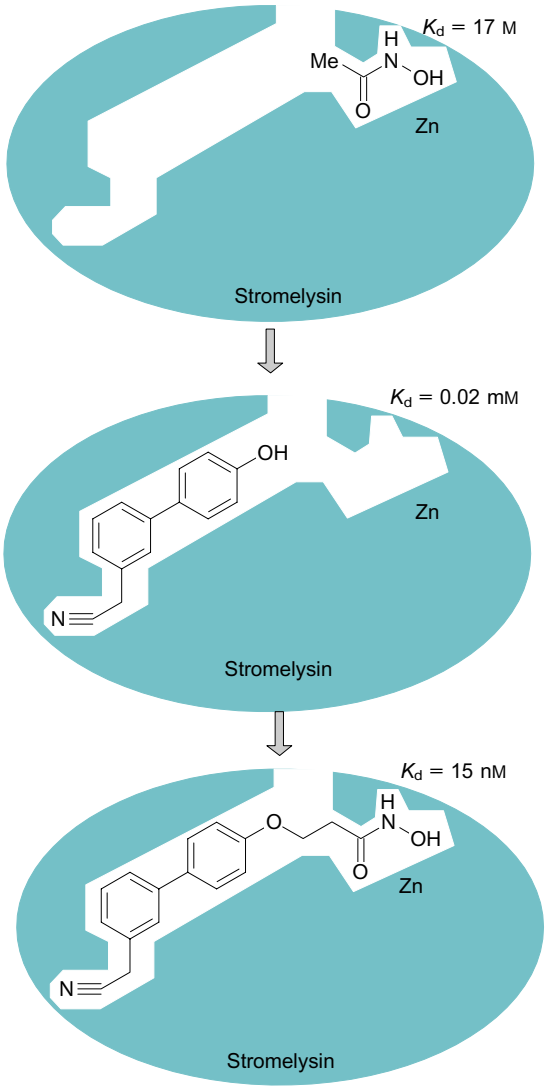


Figure 20. Design of efficient inhibitor of matrix zinc-dependent metalloprotease stromelysin-3 using the SAR by NMR method.[124]

NMR titration experiments with relatively small target proteins are most often carried out using 2D $^{15}N-^{1}H$ HSQC spectroscopy or a faster version of this heteronuclear method, SOFAST-HMQC.[125] In the case of larger proteins (> 30–40 kDa), deuteration is combined with the TROSY (Ref. 126) or CRINEPT-TROSY (Ref. 127) methods. Very large proteins (hundreds of kDa) are studied in the $^{13}C-^{1}H$ TROSY experiments using perdeuterated samples whose leucine, valine and isoleucine amino acid residues contain $^{13}CH_3$ groups.[128] Here, selective labelling is achieved by adding one or more selectively labelled amino acid biosynthesis intermediates, *e.g.*, [3-^{13}C]-pyruvate, [4-^{13}C]-α-ketobutyrate or [4,4′-^{13}C]-α-ketoisovalerate to the expression medium.[84] This approach made it possible to assign methyl groups of the key hydrophobic residues of thermophilic archaea proteasome[129] and more recently to design a new proteasome inhibitor, the drug bortezomib[130] approved by the US Food and Drug Administration.

The SAR by NMR method is highly informative and has been used to design a rather large number of novel biologically active compounds.[131] However, the approach is expensive since it involves the preparation of ^{15}N- and ^{13}C-labelled protein samples and signal assigmnent in the NMR spectra of proteins.

4.2.3. Application of proteins selectively labelled with stable isotopes

Recently, new versions of the SAR by NMR method based on the use of proteins selectively labelled with the stable isotopes ^{13}C, ^{15}N and ^{19}F were proposed. One of them, RAMPED-UP NMR, assumes operation with proteins in which certain amino acid residues are labelled with ^{13}C and(or) ^{15}N isotopes.[132] Such proteins are obtained by, *e.g.*, expression of a protein in a cell-free system containing only particular types of ^{13}C- and ^{15}N-enriched amino acid residues.[133] Combinatorial selective labelling of proteins is also suitable.[134] In this case the HSQC or TROSY spectra of the protein are significantly simplified and one can simultaneously study a mixture of proteins. It is possible not only to detect selective binding of a compound to a certain protein from the mixture in hand, but also to monitor protein – protein interactions.[132] It should be noted that efficient application of the RAMPED-UP NMR method is only possible having signals of all proteins assigned and the ligand binding sites determined.

Fluorine-19 atoms are highly sensitive sources of information obtained using NMR spectroscopy methods. Rather efficient methods for introducing ^{19}F-labelled amino acids (mainly aromatic ones) into protein molecules have been developed.[135, 136] In this case even 1D NMR spectroscopy provides valuable information for monitoring the interaction between ^{19}F-labelled proteins and low-molecular-mass ligands. The PrOF NMR method[135, 137] was developed for screening ligands capable of binding to ^{19}F-labelled proteins. It was tested using, *e.g.*, bromodomains Brd4 and BrdT;[137] transcription factor binding domain KIX;[138] the SPRY domain of SPSB2 protein regulating the activity of inducible NO synthase,[139] and AMA1 protein, a target of antimalarial drugs.[140] In all cases, proteins containing either 5-fluorotryptophan or 3-fluorotyrosine were obtained. The method is characterized by high sensitivity and performance (screening of hundreds compounds within a few hours) and allows one to determine the lead binding constants.

4.2.4. Application of ligands containing spin labels or paramagnetic ions

Yet another extended version of the SAR by NMR method involves the use of ligands containing spin labels[141] or lanthanide ions.[142] A considerable paramagnetic contribution to the transverse relaxation rate as well as the effect of pseudocontact paramagnetic shift observed for Tb, Dy, Ho, Eu and Yb atoms and much exceeding the effects of ring currents enable the detection of signals from protein domains surrounding labelled ligands. For instance, the binding site of anesthetic agent sevoflurane to calmodulin was located[143] using ligands attached to a chelate complex of Dy^{3+}. It was reported that this approach may appear to be efficient for locating the binding sites of various ligands interacting with low affinity target proteins.

4.2.5. Comparison of the NMR screening methods based on monitoring the target signals

Protein – ligand interactions are monitored using highly sensitive heteronuclear spectroscopy methods, namely, HSQC for relatively small and TROSY for large protein molecules. Practical implementation of both methods requires isotopic labelling of proteins with ^{15}N, ^{13}C and ^{2}D (for large proteins), information on the assignment of protein signals in the NMR spectra and on the structure of the protein target. Methods based on monitoring of protein signals are highly informative. They enable mapping of the ligand binding site and provide direct information for subsequent structure optimization. Selected parameters of these NMR screening methods are listed in Table 6.

As to the drawbacks, the SAR by NMR method is expensive since it involves the use of relatively high concentrations of labelled proteins and limits the molecular mass of the biomolecules under study. Nevertheless, progress in NMR spectroscopy and methods of isotope labelling of proteins[81] makes it possible to study large macromolecules. For instance, the SAR by NMR experiments using ^{15}N-enriched proteins can be used to study systems with molecular masses up to 25 – 30 kDa and this limit can be increased to 100 – 150 kDa when using deuterated proteins and the TROSY (Ref. 144) and CRINEPT (Ref. 127) methods.[145] The use of deuterated proteins selectively labelled with ^{13}C at methyl groups of isoleucine, leucine, valine and alanine residues[129] enables further increase in the molecular mass of the objects of interest. An example, the interaction of 20S-proteasome (750 kDa) with a series of low-molecular-mass compounds including bortezomib, the only approved proteasome inhibitor, was studied[130] using NMR spectroscopy methods.

4.3. NMR screening methods based on detection of the ligand signals

Compared to other biophysical methods, NMR spectroscopy gives the best results in studies of the interaction between proteins and compounds with relatively low binding constants (see above). In this case one deals with equilibrium between the free and protein-bound states of the ligand. If the ligand is present in excess, the chemical shifts of its signals are almost unaffected by the interaction with a protein. Therefore, monitoring of the protein – ligand interaction requires information on other parameters, namely, the rate of magnetization transfer, the sign of the NOE, the translational diffusion coefficient of ligand molecules and the transverse relaxation rate (see Table 5). These parameters are significantly different for low-molecular-mass compounds and large biomolecules. When bound to a protein, the ligand acquires the properties of a high-molecular-mass compound. Clearly, under equilibrium between the free and bound states of the ligand the value of an observed parameter (A_{obs}) in the NMR spectrum of the ligand is given by

$$A_{obs} = A_f(1 - p_b) + A_b p_b \quad (30)$$

Here A_f and A_b are the NMR spectral parameters of the free and bound forms of the ligand, respectively, and p_b is the proportion of the bound form of the ligand.

Table 6. NMR screening methods based on monitoring of signals in the NMR spectra of the target and knowledge of its 3D structure.

Method	Molecular mass of protein /kDa	Number of ligands in the mixture	Working concentration /mM		Note
			protein	ligand	
HSQC	< 30	< 5 [a]	~0.05–0.5	~0.2 –5	^{13}C and(or) ^{15}N-labelled protein
TROSY and CRINEPT	< 100	< 5 [a]	~0.05–0.5	~0.2–5	^{2}D, ^{15}N and(or) ^{13}C-labelled protein
RAMPED-UP NMR	< 30	< 3 [b]	~0.1–0.5	~0.5–5	^{15}N and(or) ^{13}C selectiely labelled protein
PrOF NMR	< 30	< 10	~0.05–0.5	~0.5–5	Protein containing ^{19}F-modified amino acid residues (5-F-Trp, 3-F-Tyr)

Note. Numerical values were taken from Refs 29, 37, 38, 132 and 135. [a] Typically, one ligand is studied in a particular experiment. [b] Usually, one ligand and a number of proteins are studied in a particular experiment.

Even if the p value is small, but the property of the ligand differs significantly upon binding to the biomolecule, the A_{obs} value will differ appreciably from A_f. It follows that by detecting the properties of the ligand one can reliably identify its interaction with the biomolecule. In the text below we describe main methods for monitoring the protein–ligand interaction *via* detection of the ligand signals.

4.3.1. The STD and WaterLOGSY methods

The difference spectroscopy methods STD and WaterLOGSY are based on the high rate of spin diffusion in large molecules; they are suitable for studying the interaction of mixtures containing two to ten ligands and a target. These experiments require a proper choice of compounds in order to identify resonances of all ligands, *i.e.*, their signals should not overlap in the 1D spectrum.

In the STD method [23, 146] one or more resonance frequencies of a protein are selectively saturated using a series of frequency-selective pulses (Fig. 21 *a*), while the resonance lines of ligands remain untouched. This can be done in almost all cases since the range of the chemical shifts for proteins is much wider than for ligands and protein resonance(s) in the upfield ($\delta < 0$) or downfield ($\delta > 10$) regions can be found with ease.

Saturation of a protein signal over a period of 2–5 s causes rapid transfer of the magnetization over the entire biomolecule by the spin diffusion mechanism until the protein-bound ligand is reached (Fig. 21 *b*). As a result, the intensities of signals of the ligand that appeared repeatedly at the binding site of the protein during the long saturation time decrease. This is also favoured by slow relaxation of magnetic nuclei of the free ligand that leaves the complex with the protein upon dissociation. These effects are not observed for the nonbinding ligands. Reliable identification of binding is achieved using the difference spectroscopy method. To this end, the spectrum measured under saturation of a protein resonance ('on') is subtracted from the spectrum measured under irradiation of the empty region where no NMR signals are observed ('off'). Thus the difference spectrum exhibits only the ligand signals whose intensities decrease upon saturation of protein resonances (Fig. 21 *c*). The protein signals are suppressed using the $T_{1\rho}$-filter, *i.e.*, a low-intensity rf pulse that causes the magnetization to hold in the perpendicular plane during

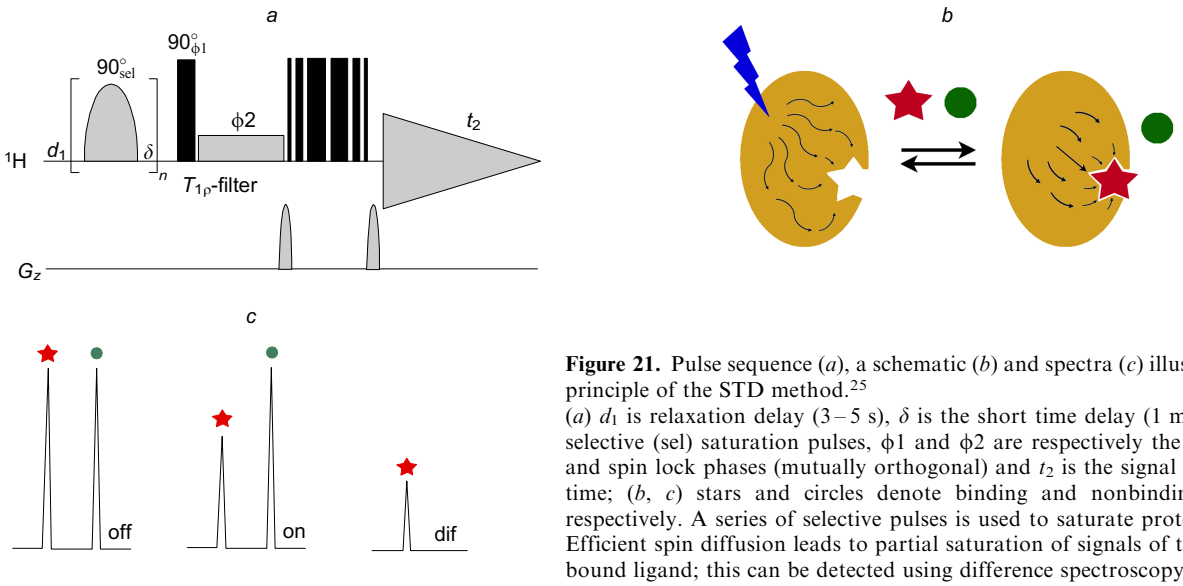


Figure 21. Pulse sequence (*a*), a schematic (*b*) and spectra (*c*) illustrating the principle of the STD method.[25]
(*a*) d_1 is relaxation delay (3–5 s), δ is the short time delay (1 ms) between selective (sel) saturation pulses, ϕ1 and ϕ2 are respectively the read pulse and spin lock phases (mutually orthogonal) and t_2 is the signal acquisition time; (*b*, *c*) stars and circles denote binding and nonbinding ligands, respectively. A series of selective pulses is used to saturate protein signals. Efficient spin diffusion leads to partial saturation of signals of the protein-bound ligand; this can be detected using difference spectroscopy method.

the time interval necessary for relaxation of nuclei producing broad signals.

The STD method is of paramount importance to the search for compounds capable of binding to targets.[75,147–149] Wide use of this approach is accompanied by its modifications. In particular, optimal methods for selective saturation of target signals were developed[136,150] to improve the sensitivity of the STD method. The resolution can be enhanced using 2D methods STD-TOCSY and STD-HSQC.[151,152] Protocols for determining the ligand binding constants from the STD data were developed and the applicability range of the STD method for measurements of binding constants K_a was established.[153–155] A heteronuclear version of the STD method with saturation of protein protons followed by magnetization transfer to ^{19}F nuclei of the bound ligand was elaborated.[156] The technique offers a rather high resolution and sensitivity in studies of weak specific binding of carbohydrates and their mimetics to lectins.

The evolution of the STD experiments continued with the development of the saturation transfer double difference (STDD) spectroscopy.[157] This approach allows one to identify the ligand signals in complex mixtures, *e.g.*, cell suspensions or cell lysates containing large amounts of small molecules. Two STD spectra of the mixture of interest are measured in the STDD experiments, one in the absence and one in the presence of a ligand. Then, one spectrum is subtracted from the other. This procedure was used in a study of the binding of the cyclic pentapeptide cycloRGDfV to the glycoprotein αIIbβ3 of intact blood platelets.[157] A similar method was used to identify ligands that bind to the protein regulating the modification of sialic acid residues on the surface of cell membrane.[158] These STDD experiments were carried out using a Sepharose-immobilized protein.

Yet another difference NMR spectroscopy method widely used in NMR screening experiments is WaterLOGSY (Fig. 22).[159] It enables highly sensitive detection of the interaction of proteins with low-molecular-mass compounds and does not require pre-saturation of protein signals.

In WaterLOGSY experiments, the main source of information on the change in the molecular mass of the studied ligand upon its binding to the target protein is the rate of cross-relaxation of the ligand protons with water molecules. As mentioned above, the signs of the cross-relaxation rates responsible for the magnitude of the NOE are different for low-molecular-mass compounds and proteins (see Fig. 16 and Table 5). The WaterLOGSY experiment involves inversion of the water proton magnetization that is transferred to the ligand protons by the cross-relaxation or chemical exchange mechanism during the mixing time (see Fig. 22 *b*). By subtracting this spectrum from the spectrum with the equilibrium signal of water protons it is possible to differentiate between the binding and nonbinding ligands from the signs of their signals in the difference spectrum (see Fig. 22 *c*). The subtraction is performed automatically during the acqusition of the WaterLOGSY spectrum. Like in the STD experiments, a mixture of ligands in the presence of the target protein is studied. Typical concentrations are a few tens of μM for the protein and hundreds μM for ligands (Fig. 23).

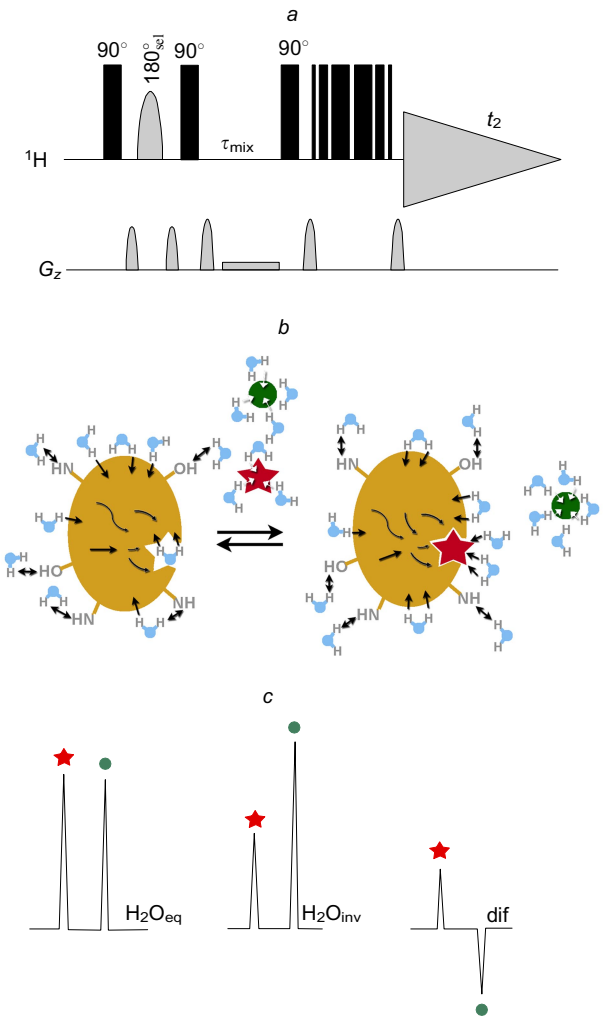


Figure 22. Pulse sequence (*a*), a schematic (*b*) and spectra (*c*) illustrating the principle of the WaterLOGSY method.[25]
Stars and circles denote binding and nonbinding ligands, respectively. Application of selective pulse (sel) causes inversion of the water signal. During the mixing time (τ_{mix}) inverted magnetization is transferred to the protein nuclei and then to the nuclei of the bound ligand through chemical exchange and cross-relaxation (NOE). Analogous cross-relaxation effects for nonbinding low-molecular-mass compounds have opposite sign. Scans without and with inversion of the water signal (H_2O_{eq} and H_2O_{inv}, respectively) are summed and subtracted. Eventually, binding and nonbinding ligands differ in the sign of their signals in the difference spectrum (dif).

Efficient use of the WaterLOGSY pulse sequence requires that the magnetization of water protons have the time to relax. As a consequence, a long relaxation delay before application of the pulse sequence is needed. To improve the water polarization recovery, a polarization optimized PO-WaterLOGSY NMR experiment was proposed.[160] It allows one to considerably reduce the relaxation delay and thus the duration of the pulsed experiments (by a factor of five to six). Yet another modification of the method is the aroma-WaterLOGSY technique[161] based on selective excitation of signals of aromatic protons only. On the one hand, this precludes identification of ligand binding from resonances of the

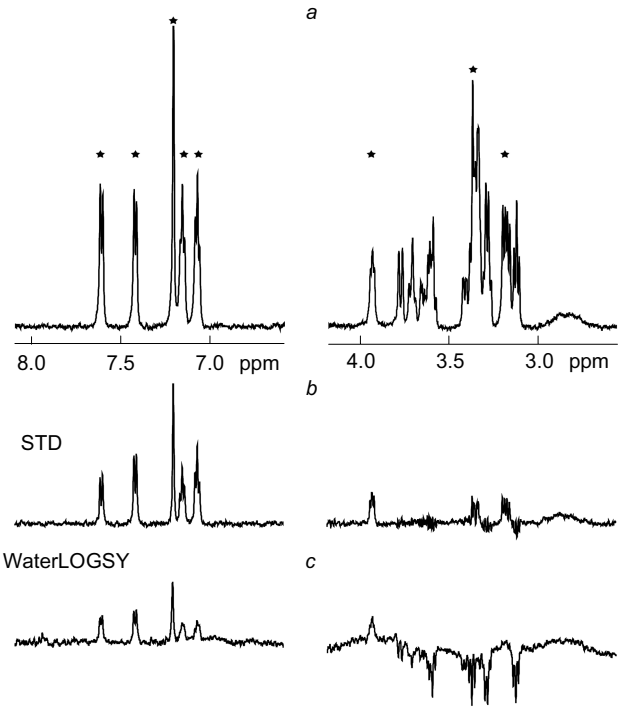


Figure 23. Examples of STD and WaterLOGSY spectra of a model mixture containing L-tryptophan (Trp, 3.2 mM) and glucose (Glc, 3.2 mM) in the presence human serum albumin (HSA, 62 μM). The protein : ligand mole ratio is 1 : 50 for both ligands. Blank 1D spectrum of the mixture (*a*); STD spectrum (*b*) and WaterLOGSY spectrum (*c*). Tryptophan can bind to HSA. Its signals (asterisked in spectrum *a*) are clearly seen in the STD spectrum while the signals of glucose (nonbinding ligand) are very weak. In the WaterLOGSY spectrum, these groups of signals have different signs.

ligand aliphatic protons. On the other hand, signals of aromatic protons are usually well-resolved, rather narrow and intense. Besides, aromatic fragments are most often deeply buried into the binding site of the biomolecule and thus the most pronounced effects in the WaterLOGSY NMR spectra are observed for just these resonance lines. The aroma-WaterLOGSY NMR approach also allows one to utilize proton-containing buffers and additives (TRIS, HEPES, glycerol, DTT, *etc*.) that are of limited use in conventional NMR experiments.

The sensitivities of the STD and WaterLOGSY methods were compared for three model ligand–protein systems, namely, ketoprofen–bovine serum albumin (69 kDa), *tert*-butylhydroquinone–influenza hemagglutinin (~170 kDa) and chloramphenicol–ribosome (70S, a few MDa).[162] The WaterLOGSY method was found to be more sensitive than the STD method for all systems studied, *viz*., by a factor of about three for relatively small albumin and by a factor of nearly 16 for the larger systems. The reason is that inversion of the water signal in the WaterLOGSY experiment simultaneously influences numerous protons of the biomolecule that act as sources of cross-relaxation with ligand protons (*cf*. saturation of only one or a few protons in the STD experiment). At the same time the attainment of optimum results using the WaterLOGSY method requires a more thorough choice of parameters (*e.g*., mixing time and relaxation delay).

A heteronuclear ^{19}F WaterLOGSY method is similar to the heteronuclear ^{19}F STD modification (see above) and can be used to detect the binding of fluoro-containing ligands to biomolecules.[163]

The STD and WaterLOGSY approaches are suitable not only for detection of ligand–protein binding, but also for determination of the binding constant K_a. This requires a series of measurements at different ligand : protein mole ratios. Most often these experiments are carried out to study an individual compound rather than a mixture of ligands. In the WaterLOGSY experiments the intensity of the *i*th signal of the ligand, I_i, depends on the efficiency of magnetization transfer from water protons to protons of the free (L) and protein-bound (PL) forms of the ligand and, to a first approximation, is given by

$$I_i \propto \sigma_{PL}[PL] + \sigma_L[L] \quad (31)$$

$$\sigma_{PL} = \sum_j \sigma_{ij}^{bound} + \sum_k \sigma_{ik}^{bound} + \sum_w \sigma_{iw}^{bound} \quad (31a)$$

$$\sigma_L = \sum_j \sigma_{ij}^{free} + \sum_w \sigma_{iw}^{free} \quad (31b)$$

Here σ_{PL} and σ_L are the rates of magnetization transfer from water protons to protons of the bound and free ligand, respectively; the subscript j denotes ligand exchangeable protons (they can exchange with protons from water molecules, OH and NH groups, *etc*.), the subscript k denotes the protein protons near the ligand in the complex being formed; and the subscript w refers to water protons near the ligand.[159] The contribution of the free ligand form to the intensity of the *i*th signal in the WaterLOGSY spectrum ($I_{0i} = \sigma_L[L]$) can be measured in a blank experiment in the absence of the protein. In this case, the intensity difference of the *i*th signal of the ligand (ΔI_i) in the WaterLOGSY experiments carried out in the presence and in the absence of the protein is given by the simple equation

$$\Delta I_i = I_i - I_{0i} = \sigma_{PL}[PL] \quad (32)$$

Expression (32) can be used along with Eqn (13) to determine the dissociation constants. Changes in the intensity of a signal of the ligand bound to *C. freundii* methionine-γ-lyase are presented in Fig. 24.

A correct determination of the constant K_d in the WaterLOGSY experiments requires a thorough choice of experimental conditions. An important role is played by the concentration of the protein, which should be minimum but sufficient to provide for reliable measurements of signal intensities in the NMR spectra. An increase in the protein concentration leads to overestimation of the dissociation constants.[165, 166]

The STD and WaterLOGSY methods and their numerous modifications are the main tools of NMR screening.

4.3.2. Methods based on detection of ^{19}F signals

Recently, there has been rapid progress in ^{19}F NMR methods since they are characterized by amazingly high

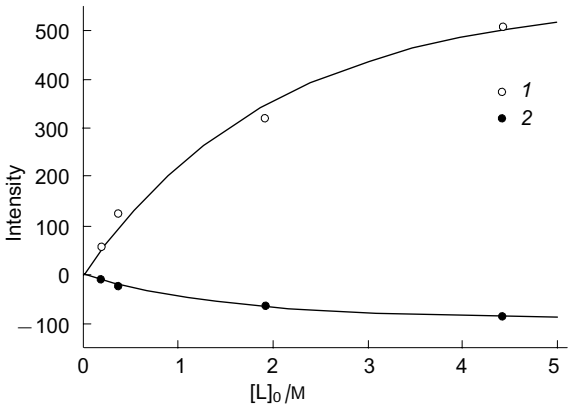


Figure 24. Integrated intensity of the ^{1}H signal of the ligand bound to *C. freundii* methionine-γ-lyase in the WaterLOGSY spectra plotted *vs.* ligand concentration. Curve *1* corresponds to the experiments carried out in the presence of the protein (46 μM). Curve *2* corresponds to the experiments carried out in the absence of the protein. Calculations with inclusion of changes in the intensity difference between curves *1* and *2* led to $K_d = 2.15 \pm 0.58$ mM.[164]

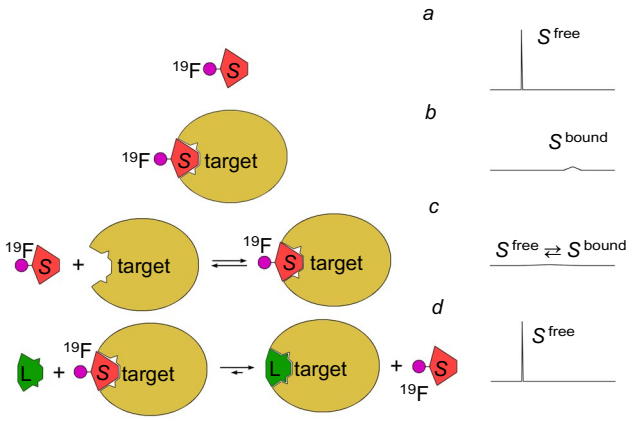


Figure 25. The FAXS and n-FABS methods: a schematic.[167] ^{19}F NMR spectrum of free low-molecular-mass (*S*, 'spy') compound *S* (*a*); ^{19}F NMR spectrum of bound spy compound (*b*); broadening to disappearance of the ^{19}F signal at equilibrium between the bound and free compound *S* (*c*); adding a high-affinity binding ligand L causes displacement of molecule *S* from the complex and re-appearance of the signal of free compound *S* in the ^{19}F NMR spectrum (*d*).

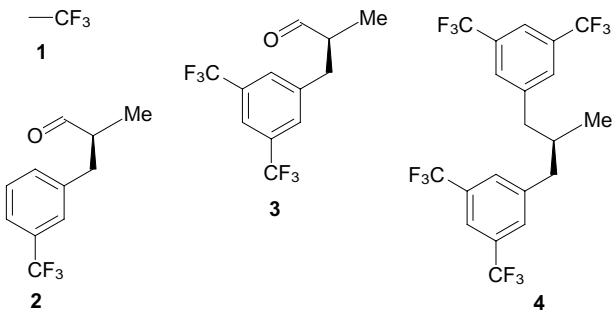
sensitivity comparable with that of fluorescence methods for monitoring protein – ligand interactions. The reasons are as follows. First, the ^{19}F isotope is almost absent in natural biological objects and thus the ^{19}F NMR spectra are rather simple. Next, the gyromagnetic ratio of ^{19}F nucleus is comparable with that of proton and thus ensures a close sensitivity. Third, the ^{19}F resonances are observed in a very wide range of chemical shift values, so even minor changes in the environment of a ^{19}F nucleus cause remarkable change in the position of the resonance line. Finally, the relaxation properties of ^{19}F nucleus are suitable for monitoring the interaction of ^{19}F-labelled ligands with targets. A research group headed by Dalvit [167, 168] developed two methods, FAXS and n-FABS, that utilize unique properties of fluorine nuclei. Both methods are based on the choice of a small ligand (often, oligopeptide) containing fluorine atoms that can bind to the target under study (Fig. 25).

A free compound *S* ('spy' compound according to Dalvit *et al.*) has a rather narrow ^{19}F signal. Therefore, its peak intensity is high (Fig. 25 *a*). The signal of the bound spy compound (i) changes the chemical shift and (ii) is significantly broadened (Fig. 25 *b*). The broadening is to a great extent due to high chemical shift anisotropy of ^{19}F nuclei. If the bound and free forms of compound *S* are at equilibrium, the signal in the ^{19}F NMR spectrum of *S* is broadened to so high extent that the resonance becomes almost indistinguishable (Fig. 25 *c*). This solution represents a test system in the search for compounds that can efficiently compete for the same binding site in the target. Binding of a ligand with higher affinity causes reappearance of the ^{19}F resonance line of the free compound *S* competitively displaced by the new ligand (Fig. 25 *d*). The method enables screening of multicomponent mixtures within short times. The sensitivity of the FAXS method using compounds containing a single fluorine atom is comparable with that of the STD and WaterLOGSY methods. However, the spy compound can contain a number of chemically equivalent fluorine atoms (structures **1** – **4**), which significantly increases the sensitivity of the n-FABS method.[168 – 170]

The methods based on the detection of ^{19}F signals proved to be versatile; they are used in studies of soluble and membrane proteins in the lipid environment,[171] as well as cell lysates and intact cells.[172] In particular, the n-FABS method was employed to study living human cells expressing the membrane enzyme FAAH and identified inhibitors capable of binding to this enzyme in a wide range of affinity values.[172]

The key advantage of the FAXS and n-FABS methods is high sensitivity that allows targets to be studied at much lower concentrations. Indeed, the protein concentration > 100 μM is necessary for the SAR by NMR methods, 10 μM for the STD and WaterLOGSY methods and can be as low as a few tens of nM for the n-FABS method. A 3-FABS experiment using highly sensitive ^{19}F cryogenic probe revealed the formation of a proteolysis product of a polyfluoro-substituted peptide under the action of only 200 fM of trypsin.[168] Screening of enzyme inhibitors whose action did not lead to product formation was performed with a test system containing a fluorinated peptide and trypsin.

Structures 1 – 4

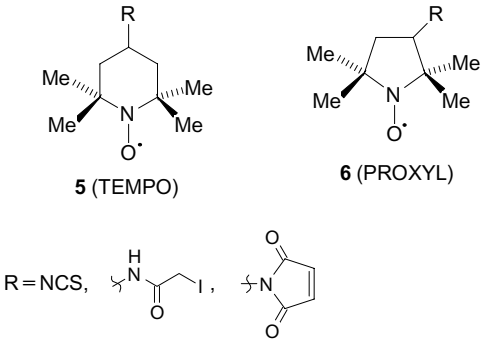


4.3.3. Methods involving spin-labelled proteins and proteins complexed the lanthanide ions

A number of methods based on modification of the target proteins with spin labels or lanthanide ions that induce paramagnetic relaxation enhancement are available. This effect significantly changes the relaxation properties of nuclei in the ligands bound to the modified protein, which allows one to identify them using NMR spectroscopy methods.

A SLAPSTIC approach[173] proposed to this end involves modification of the protein with 2,2,6,6-tetramethylpiperidin-1-oxyl (TEMPO, structure **5**) or 2,2,5,5-tetramethylpyrrolidine-1-oxyl (PROXYL, structure **6**) derivatives containing isocyanate, maleimide or 2-iodoacetamide residues as substituents. These groups can respectively bind to the side-chain of lysine, cysteine and methionine residues on the surface of the protein under mild conditions. This is a route to proteins modified with one or more stable radicals.

Structures 5, 6



The paramagnetic properties of the unpaired electron manifest themselves as a remarkable increase in the transverse relaxation rate (and, hence, broadening of signals) of atomic nuclei located within ~15–20 Å from spin label.[173, 174] This broadening can be reliably detected using a $T_{1\sigma}$- or T_2-filter that passes only narrow signals. Broad signals characterized by high transverse relaxation rates disappear within the duration of the rf pulse in the $T_{1\sigma}$-filter experiments or a series of 180-degree CPMG pulses in the T_2-filter trains. Next, by using the difference NMR spectroscopy techniques one can detect only signals of the ligand bearing the free-radical label broadened as a result of binding to the protein (Fig. 26).

It should be noted that the introduction of free-radical labels and subsequent analysis of the nuclear relaxation rates are widely used for protein structure determination. In particular, an approach involves engineering of a series of mutant forms of a membrane protein bearing cysteine residues attached to particular atoms and subsequent modification of these residues with paramagnetic labels.[175, 176] Amino acid residues surrounding the modified cysteine can be identified from the changes in the nuclear relaxation rates of ^{1}H, ^{15}N and ^{13}C of the protein resonances. This method was used to determine the structures of certain protein–RNA complexes.[177, 178] Often, the role of paramagnetic label is played by complexes with paramagnetic metal ions (*e.g.*, Mn^{2+}, Cu^{2+}, Gd^{3+}) rather than nitroxides.[178]

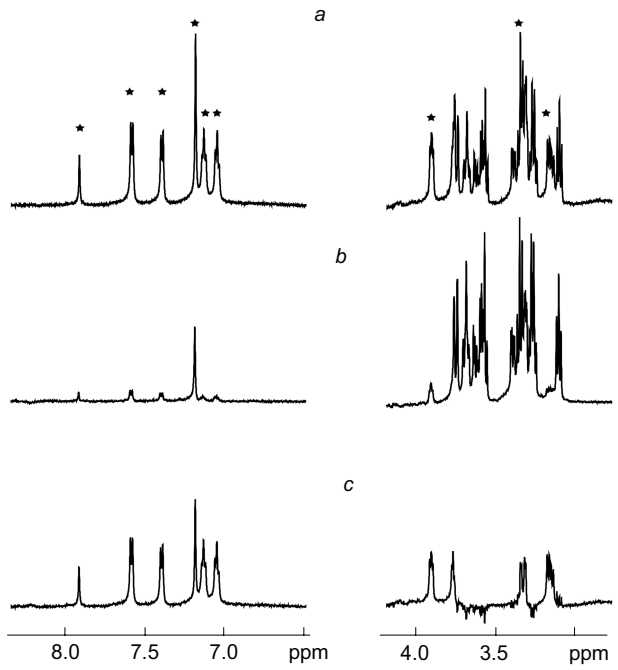


Figure 26. Results of SLAPSTIC experiment using a model mixture containing tryptophan (Trp, 3.2 mM) and glucose (Glc, 3.2 mM) in the presence of 2-iodoacetamide-modified human serum albumin (HSA, 62 μM). The protein : ligand mole ratio is 1 : 50 for both ligands. Blank 1D spectrum of the mixture (*a*); T_2-filtered spectrum (200 spin echo repetitions within 204 ms) (*b*) and the difference spectrum (*c*). The difference spectrum exhibits signals of tryptophan bound to HSA (asterisked in spectrum *a*) and no signals of glucose (nonbinding ligand).

4.3.4. Comparison of NMR methods based on detection of ligand signals

NMR methods based on the detection of the ligand signals and used to identify ligand capable of binding to the target biomacromolecules (Table 7) are much more sensitive than methods based on the detection of the biomacromolecule signals (see Table 6). Yet another distinction between these methods is the possibility to detect the binding event using a mixture of ligands; this significantly improves the performance of the screening procedure. Besides, methods with detection of the ligand signals have no limitation on the size of the biomolecule target and the observed effects are even more pronounced in the case of large biomolecules. The STD, WaterLOGSY and SLAPSTIC methods are only applicable to weakly binding ligands undergoing fast exchange between the free and bound forms. No effects of binding to the target biomolecule will be observed in the case of slow exchange between these forms. Contrary to this, the FAXS and n-FABS methods are applicable under strong binding conditions since the ^{19}F-containing 'spy' molecule can be competitively displaced from the binding site by a certain ligand only in this case.

The TINS method was proposed for screening of ligands that can bind to Sepharose-immobilized targets.[179] The ligands capable of binding are identified from changes in the ligand resonance linewidth in the absence and in the presence of Sepharose resin used as support. The technique is not highly sensitive but allows the same target sample to be reused in a large series of

Table 7. NMR screening methods based on the observation of spectral parameters of ligands.

Method	Number of ligands in the mixture	Working concentration		Note
		protein /μM	ligand /mM	
STD	2 – 20	20 – 200	0.1 – 5	One of the most widely used methods; applicable in the case of weakly binding ligands
WaterLOGSY	2 – 20	20 – 200	0.1 – 5	More sensitive method compared to STD; applicable in the case of weakly binding ligands
SLAPSTIC	2 – 10	20 – 200	0.1 – 5	Highly sensitive method that requires modified proteins; applicable in the case of weakly binding ligands
FAXS, *n*-FABS	5 – 20	10^{-3} – 50	5×10^{-2} – 1	Most highly sensitive methods; applicable to detect strongly binding ligands
DOSY	2 – 10	100 – 250	0.2 – 1	Method based on the decrease in translational diffusion rate of the ligand upon binding to the target
TINS	2 – 10	250	0.2 – 0.5	Search for ligands binding to targets immobilized on Sepharose support; one sample of a target protein can be repeatedly used to screen a number of ligand libraries

Note. Numerical values were taken from Refs 29, 38, 120, 154, 162, 168, 173 and 179.

experiments. Besides, the method can be implemented in flow probe NMR detection. Binding of the small ligands to high-molecular-mass targets is also identified by measuring the rate of translational diffusion using gradient DOSY NMR methods.[180, 181] The DOSY approach is based on a decrease in the diffusion coefficient of the ligand molecules bound to the large low-mobile target molecules. Being not highly sensitive, DOSY-based methods may appear to be rather informative in some cases.[117, 119, 182]

Selected characteristics of the described above NMR screening methods based on the detection of the spectral parameters of ligands are listed in Table 7.

5. NMR spectroscopy methods for lead design

The NMR screening methods described in the preceding Sections enable identification of small molecules as molecular fragments of a potential lead compound. In the next steps of lead design an increase in their affinity for the target is achieved by modifying functional groups or by linking a number of fragments into a single compound. Strictly speaking, NMR screening can be treated as a process utilizing various methods for (i) detection of specific binding of ligands to the biomacromolecular target and (ii) rational modification of the ligand structure in order to increase the ligand affinity.

5.1. NMR methods for mapping bound ligand epitopes

Methods based on the observation of NMR signals of a target, in combination with information about its three-dimensional structure, allow a rational design of the lead compound using the selected fragments. This can also be done if only the signals of small molecules are detected in the NMR experiments. In this case information on 3D structure of a receptor or its complex with a certain hit compound is also useful. However, necessary information not always can be obtained by X-ray crystallography or NMR spectroscopy methods owing to sophisticated experiments and high labour costs.

The NMR spectroscopy methods based on the detection of ligand signals include INPHARMA,[183] SAR by ILOE,[184, 185] SALMON,[186] and STD.[187] They enable mapping ligand epitopes‡, *i.e.*, determining the ligand fragments immediately involved in the interaction with the biomolecule. Besides, they allow one to obtain information on mutual orientation of various structural fragments in adjacent binding sites of the target. Similarly to the STD method (see above) these approaches are efficient in the case of ligands with fast exchange between the free and receptor-bound forms. Experiments are carried out using biomolecular targets with molecular masses >20 kDa provided that the ligand is present in a 10 – 100-fold excess. Typically, the concentration of the target is in the range of 10 – 25 μM.

In addition to the methods based on structural information on the ligand binding site on the protein surface, strategies based on rational modification of the ligand structure are available, namely, SHAPES, NMR-SOLVE and ALARN-NMR. They allow compounds with high affinity for the target protein cost-efficient design of.

5.1.1. The INPHARMA method

The INPHARMA method provides information on the mutual orientation of two ligands competitively binding in the same receptor pocket.[183] A sample containing a macromolecular target and two ligands (we denote them A and B) is exposed to a conventional proton 2D NOESY experiment. During the mixing time in NOESY experiment, protons of the ligand A bound to the receptor transfer their magnetization to the target (Fig. 27). As ligand A leaves the receptor binding pocket, ligand B binds to the receptor and the magnetization of A-protons is transferred from the target to B-protons by the spin diffusion mechanism. As a result the NOESY spectrum will exhibit cross-peaks between A-protons and

‡ The term 'ligand epitope' was introduced by analogy with the term 'antigen epitope' denoting the part of an antigen that is recognized by macromolecules of the immune system.

B-protons since both ligands are located in the same binding pocket in the case of successive binding to the receptor.

Exact ligand/target mole ratios and the NOESY mixing time (t_m) should be chosen experimentally depending on a particular system under study. For large targets (molecular mass >400 kDa) one can use $t_m < 100$ ms, while experiments with small receptors (molecular mass ~20 kDa) require long mixing time (>500 ms).[188] The exchange rates of both ligands in the receptor binding pocket should be comparable since the most pronounced INPHARMA effects are observed at equal ligand dissociation constants k_{off}.[188]

When interpreting the INPHARMA effects, one should take account of the fact that the magnetization can be transferred through the receptor spin system to a distant-from-the-ligand-proton region of the binding site by the spin diffusion mechanism.[183] Therefore, the cross-peak between A-protons and B-protons might be an artifact not necessarily indicating spatial proximity of these atoms to the same receptor site. Exact pathway of magnetization transfer in the protein–ligand complex can be determined using the full relaxation matrix approach.[189] This requires knowledge of the structures of the complexes of both ligands and allows the INPHARMA method to be used for selecting correct ligand conformations among the models obtained by molecular docking.[2, 188, 190]

The INPHARMA method was for the first time applied in the binding studies of epothilone A and baccatin III to tubulin.[183, 191] More recently, crystallographic conformations of competitive ligands in a model system with protein kinase A as target were reproduced using the INPHARMA data.[2] The method was successfully used in many other studies of protein–ligand systems.[192–195] In particular, rational optimization of the inhibitor of platelet receptor glycoprotein VI (GPVI) revealed[192] that the angiotensin II receptor antagonist Losartan and a 12 amino acid peptide (pep-10L peptide) compete for binding to GPVI; however, the peptide affinity appeared to be higher. The peptide and Losartan fragments bound to the same receptor pocket were determined and their mutual orientation was established by the INPHARMA method. Based on these data and on information on the structure of the bound peptide form the Losartan molecule was modified to reproduce key interactions of the pep-10L pharmacophore. Optimization led to a novel inhibitor of GPVI whose affinity was comparable with that of the peptide pep-10L.

5.1.2. The SAR by ILOE method

The SAR by ILOE method allows determining mutual orientation of two ligands bound in the adjacent receptor pockets.[185, 196, 197] Like INPHARMA, this approach is based on 2D NOESY NMR experiment. The spectrum is measured for a sample containing a target and a number of ligands. If ligands A and B simultaneously bind to the target in the adjacent binding pockets, the NOESY spectrum should exhibit cross-peaks between proximate protons of these ligands (Fig. 28).

The NOESY mixing time is usually varied in the range of 600–800 ms.[198] It should be noted that the intermolecular cross-peaks in the ILOE spectrum appear as a result of direct dipole–dipole interaction between protons from two ligands separated by distances shorter than 6 Å. In contrast to the INPHARMA method, spin diffusion is the source of artifacts for the SAR by ILOE experiment; therefore, it is recommended to use perdeuterated target samples.[197]

The SAR by ILOE method was successfully employed for rational design of protein inhibitor Bid involved in the control of the apoptotic signal cascade.[185, 199] A series of NOESY experiments using samples containing a target protein and mixtures of 6 to 24 ligands selected from a library consisting of 300 various structural fragments revealed a number of binding ligands and those which had ILOE cross-peaks were detected. Based on analysis of the changes in the receptor ^1H and ^{15}N chemical shifts during titration with molecular fragments, the binding sites of the hit ligands in the hydrophobic groove on the surface of the protein were mapped. The structure of the optimal linker for connecting two hit fragments was chosen using the results of molecular docking and ILOE data.

5.1.3. Application of the STD method to epitope mapping

The STD method (see above) can be used to map the ligand epitope and to reveal functional groups playing a

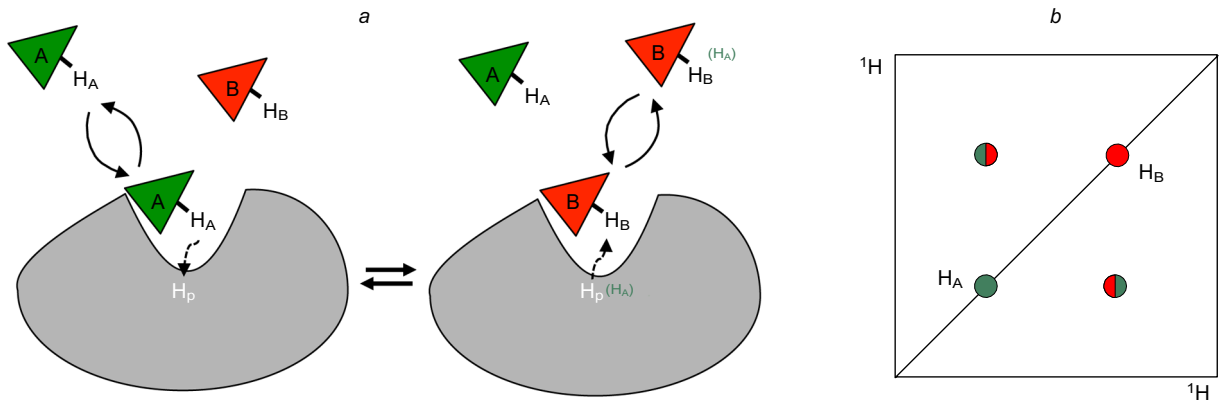


Figure 27. Schematic representation of the principle of the INPHARMA method (*a*) and NOESY spectrum exhibiting cross-peaks between protons of ligands A and B due to magnetization transfer through the receptor (*b*).[183]

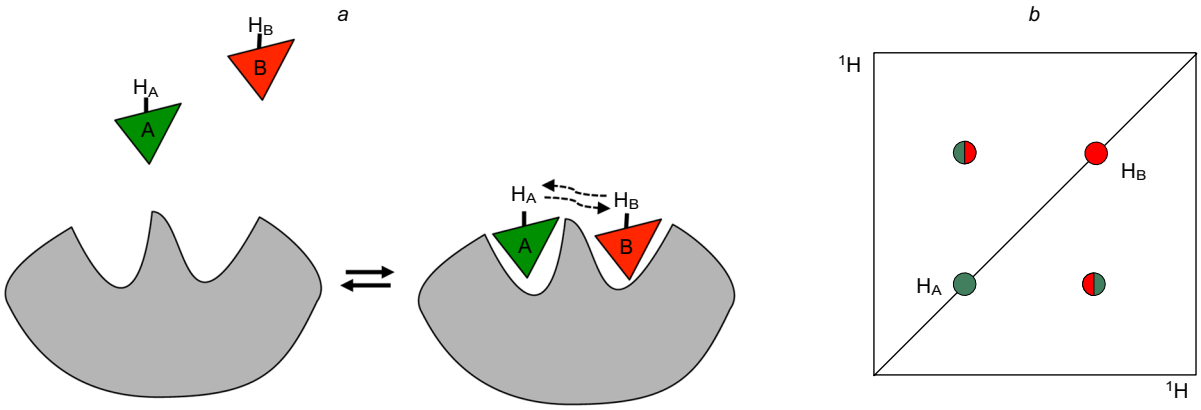


Figure 28. Schematic representation of the principle of the SAR by ILOE method (*a*) and NOESY spectrum exhibiting cross-peaks between protons of ligands A and B due to their direct interaction in the protein – ligand complex (*b*).[185]

key role in the interaction with the receptor.[187] Saturation transfer from the target to the ligand occurs at different rates depending on the distance between the ligand protons and receptor protons. Therefore, in the 1D STD spectrum, peaks corresponding to the protons in the vicinity of the target will be more intense than peaks from distant protons. To observe this effect, the residence time of the ligand in the receptor binding pocket should be short (dissociation rate should be of the same order of magnitude or higher than the rate of saturation transfer from the receptor to the ligand); otherwise, all STD peaks will look identically as a result of spin diffusion.[187] The design and analysis of the STD experiment also require taking account of longitudinal relaxation of the ligand in the presence of receptor. If the ligand proton times T_1 differ appreciably from one another, being shorter than the saturation time of the spin system of the target, the magnetization transfer by the longitudinal relaxation mechanism begins to compete with cross-relaxation between the target and ligand and contributes largely to the observed STD effects.[200]

To determine the absolute magnitude of the STD effect, the STD amplification factor (f_{STD})[187] is calculated:

$$f_{STD} = \frac{I_0 - I_{STD}}{I_0} n \qquad (33)$$

where I_0 is the intensity of a signal in the 1D reference spectrum of the ligand (no saturation of the receptor), I_{STD} is the intensity of a signal in the STD spectrum with saturation of the receptor and n is the molar excess of the ligand relative to the receptor. As the ligand excess increases, the receptor pocket is saturated and the amplification factor reaches a plateau.[187] Like in the case of the INPHARMA experiment, quantitative interpretation of the amplification factors requires full relaxation matrix calculations for the receptor – ligand complex. The results obtained can be used to validate the results of docking.[201 – 203]

The STD mapping of ligand epitope is mainly used when studying large ligands, *viz.*, sugars, peptides, peptidomimetics or vitamins.[193, 203 – 209] The problem of signal overlap in the 1D STD spectra of large molecules can be solved using 2D experiments, *e.g.*, 2D STD TOCSY[187, 207] and 2D STD HSQC.[209]

Yet another version of the STD technique called SOS-NMR[210] is based on the use of a series of deuterated proteins and allows the ligand binding site to the biomolecule as well as ligand arrangement in the binding site to be determined from spectral data. Practical implementation of this method requires the synthesis of perdeuterated proteins containing certain types of selectively ^1H labelled amino acid residues (*e.g.*, Ile, Val, Leu, Trp); this can be done by expression of the protein in the deuterated medium in the presence of excess amount of corresponding amino acids. A series of STD experiments can be used to identify the type of amino acid residues bound to the ligand in the binding site. Also, there is no need for information on the assignment of the protein signals and on the 3D structure of the protein.

5.1.4. The SALMON method

The WaterLOGSY experiment can be used to determine the ligand fragments that are in direct contact with the solvent in receptor – ligand complex.[186] A WaterLOGSY study of the binding of tretazicar to quinone oxidoreductase 2 revealed that certain protons of the bound ligand show a positive NOE (typical of the free form of the ligand) in the presence of receptor, whereas other protons exhibit negative signals typical of the bound form.[186] A more detailed investigation showed that protons with positive NOE belong to the aziridine group oriented from the surface of the target towards the solution. To sum up, the magnetization is transferred from water molecules to ligand protons inside the receptor pocket through protein atoms, whereas protons of the ligand groups distant from the protein surface directly interact with highly mobile solvent molecules (Fig. 29). This method for mapping solvent-accessible ligand fragments was named SALMON.

5.1.5. Determination of the bound ligand conformations using the trNOE method

Transferred NOE (trNOE) experiments are carried out considering an equilibrium between free and protein-

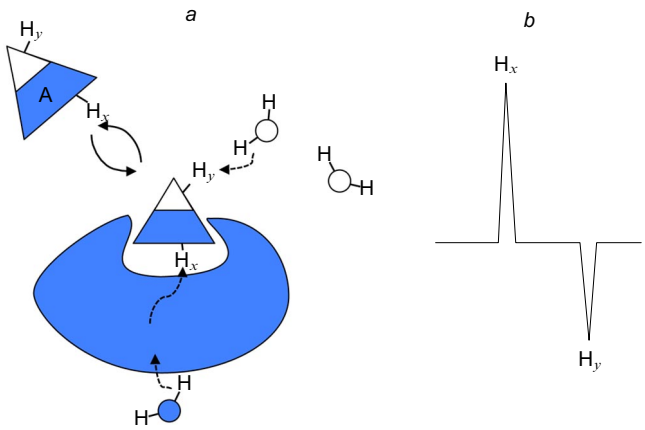


Figure 29. Schematic representation of the principle of the SALMON method (*a*) and corresponding WaterLOGSY difference spectrum of ligand (*b*).[186]
(*a*) A-ligand molecular fragments with negative NOE are given in colour, while A-ligand molecular fragments with positive NOE are unshaded.
(*b*) Ligand protons in the receptor pocket (H_x) have negative NOE and positive signals in the spectrum; protons oriented from the receptor surface towards the solution (H_y), have positive NOE and negative signals in the spectrum.

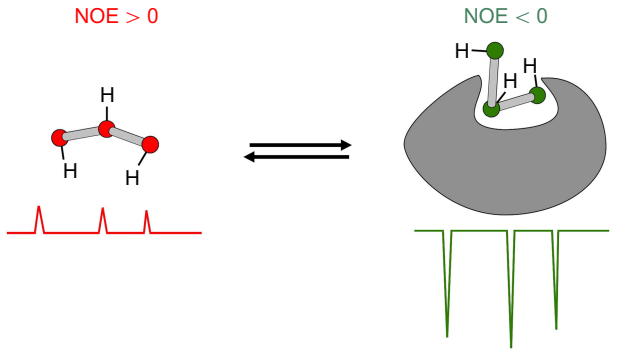


Figure 30. Schematic representation of the principle of the trNOE method.[36]
Shown is the ligand that can change its conformation when bound to the biomolecule. In the absence of a target the intensities of positive NOE between ligand protons is low. Binding to a target causes the NOE sign to change and the signal intensities in the NOESY spectra to increase.

bound states of the ligands and only under fast exchange conditions, which corresponds to observation of a single set of averaged signals in the NMR spectrum of the ligand. Typically, the method requires a rather large excess of the ligand with respect to the protein. Like in most other NMR screening methods, the observed trNOE effects are based on a considerable difference between the correlation times of rotational diffusion of the free and protein-bound forms of the ligand. For the free form, the NOE between ligand protons are positive and small, while for the bound form, they are large and negative (Fig. 30). As the magnetization did transfer to the free ligand as a result of exchange between the bound and free forms, it can be conserved for long since the relaxation rates of the free-ligand signals are low. Thus, by measuring the NOE between ligand protons in the NOESY spectra we in fact determine the conformation of the protein-bound ligand. This is of particular importance when studying conformationally flexible ligands.

The theoretical background of the trNOE approach and methods for the most accurate determination of proton–proton distances were developed in the early studies (see, *e.g.*, Ref. 211). To this end, a few measurements with different mixing times are carried out or a combination of the NOESY and ROESY methods is used.[212]

The trNOE method was used to determine the conformation of the cofactor NAD+ in the complex with yeast alcohol dehydrogenase (~150 kDa).[213] The conformations of acetylcholine analogues bound to the M2 muscarinic acetylcholine receptor (M2 mAChR) were determined by analyzing the signal intensities in their trNOE spectra.[214] A trNOE study[215] of the interaction of furylacryloyl amino acid derivatives with angiotensin-converting enzyme (130 kDa) allowed one to map the active site of the enzyme to which the substrate and inhibitor (captopril drug) bind. The structures of disaccharides in complexes with *Aleuria aurantia* agglutinin were determined by the trNOE method.[216] This technique was developed long before other NMR screening methods, but still remains in demand for studies of protein–ligand interactions.

5.2. NMR-based lead design strategies

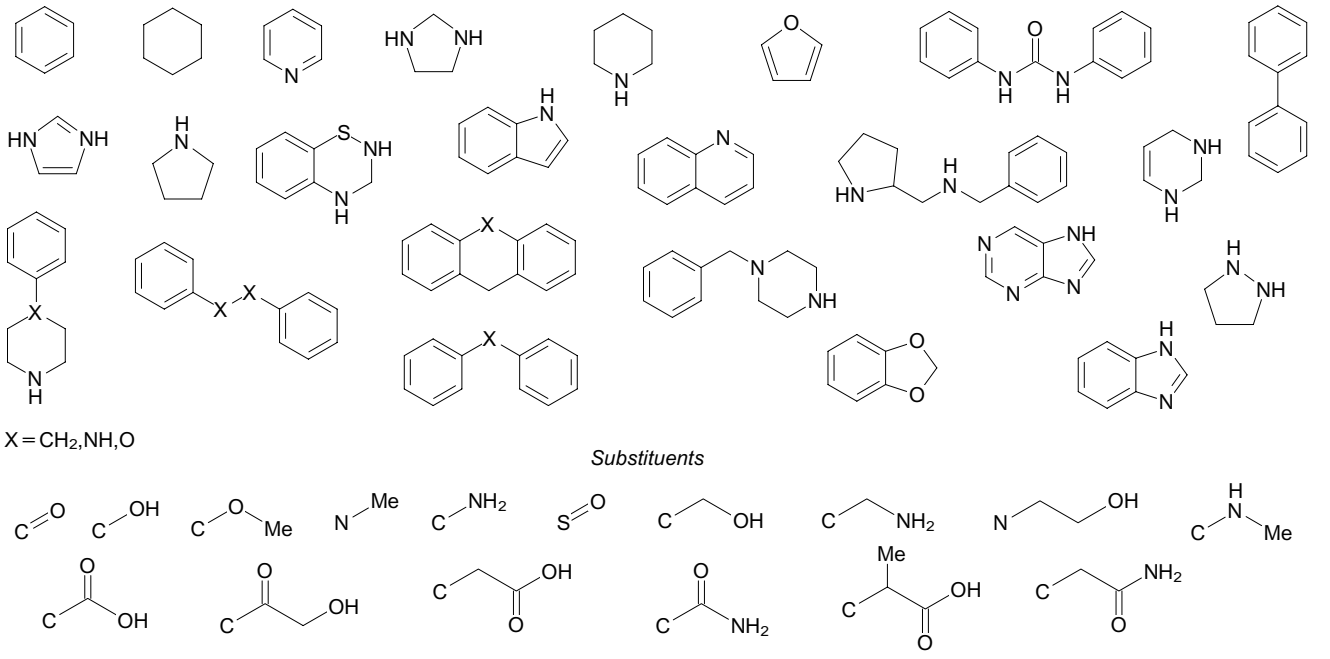
Methods for detection of the binding of molecular fragments and mapping ligand epitopes can be used to design high-affinity lead compounds interacting with targets in functionally important sites. There are various lead design strategies involving NMR spectroscopy methods and based on the detection of signals from small molecules, *viz.*, SHAPES, NMR-SOLVE, ALARM-NMR, *etc.*

5.2.1. The SHAPES strategy

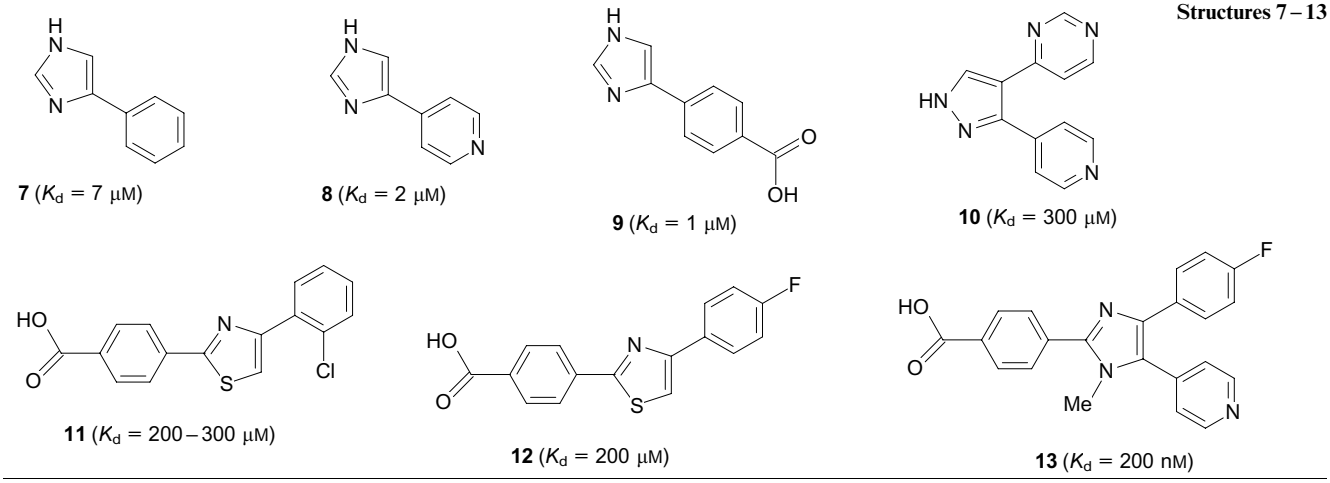
The SHAPES strategy[217] is based on the use of a relatively small but diverse library of small molecules[218] for NMR screening, which includes privileged scaffolds and key substituents. Initially, the SHAPES library included only 132 compounds. The basic molecular frameworks and key side-chain substituents used in this library of chemical compounds[218] are presented below.

The libraries for the SHAPES strategy are composed by simultaneously optimizing various factors including the cost of compounds, synthetic availability, solubility in water, the absence of signal overlap in the NMR spectra of the fragments as well as structural diversity.[217] In the first step, one or more hit compounds with the most frequently occurring molecular frameworks are selected and ranked with respect to their ability to bind to a target (including low affinity compounds). Then, the hit structure is complicated by optimizing the arrangement of substituents and(or) their structures and by gradually complicating the structure of substituents. For instance, a search for p38 MAP kinase inhibitors by NMR screening methods[217] began with imidazole as the starting point, although this compound does not bind to the enzyme of interest. Bicyclic compounds **7**–**9** obtained in the next step were found to have much higher affinity.

Structures of molecular frameworks and substituents



Structures 7 – 13



Substituents in other positions of the central ring were introduced to construct and test imidazole analogues **10** – **12**. Combining the chosen substituents in three positions of the imidazole ring led to compound **13** with K_d = 200 nM.

A SHAPES screening was used to select the structures of compounds binding with high affinity to human fatty acid binding protein (FABP-4).[24] The SHAPES library was employed to find a lead series that specifically bind to RNA target (P4 – P6 domain of the *Tetrahymena thermophila* Group I intron) using with WaterLOGSY, STD and trNOE NMR screening methods.[219] There has been progress in methods for optimization of compound libraries for NMR screening[220] and selection of compounds having minimal signal overlap.[221] A statistical analysis of the structures of small ligands using the results of NMR-based screening[222] made it possible to select privileged molecules (in particular, biphenyl derivatives) for high-affinity binding to biomolecular targets.

5.2.2. The NMR-SOLVE, NMR-DOC and ALARM-NMR methods

Screening strategies for obtaining structural information on the ligand binding sites in target proteins using relatively small anount of spectral data are available. The NMR-DOC method[223] based on a rapid protocol for obtaining of structural information by selective detection of key interactions in the binding site of the target protein. To implement this protocol, the protein is deuterated and $^{13}CH_3$ groups are introduced into particular amino acid residues (see Section 4.2.2).[128] The protocol provides structural information on protein – ligand complex even for large multimeric proteins with molecular masses >100 kDa. The NMR-DOC method was used to map the binding site of tetrameric enzyme dihydrodipicolinate reductase (DHPR, 120 kDa).[223] In the case of proteins with two closely spaced binding sites (this is typical of enzymes possessing cofactors), structural information on the first binding site can be used to design a linker reaching the second binding site.

To solve this problem, the authors of the NMR-DOC method proposed yet another strategy, NMR-SOLVE, and illustrated its potential in relation to the design of 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DOXPR, 174 kDa) inhibitor.[223] It was reported that the NMR-SOLVE method is efficient in studies of large proteins and proteins with unknown 3D structure. The NMR-SOLVE strategy is based on deuteration and selective isotopic labelling of particular amino acid residues and analysis of information on inter-ligand contacts extracted from the NOESY spectra. Structural information on metal-binding proteins can also be obtained though detection of ligand (and protein amino acid residue) line broadening upon binding of paramagnetic metal ion probes.

Screening of ligand binding to target proteins requires identification of false-positive results[224] that can be caused by not only features of the methods for detecting protein–ligand interactions (see Table 3), but also the effect of low-molecular-mass compounds on the protein, which leads to its oxidation, alkylation, aggregation, *etc.* Often, protein-related problems originate from thiol groups of cycteine residues that can be oxidized or modified. A method for reliable identification false-positive results called ALARM-NMR[225] is based on comparative detection of the interaction between the ligand under study and the protein in the absence and in the presence of a large excess of dithiothreitol (DTT) or similar reducting agents. If a highly specific ligand binding to the protein does occur, the reducing agents remain silent. However, if the ligand modifies thiol groups of the protein, the binding process is inhibited in the presence of DTT.[226]

An ALARM-NMR study[227] of a few tens of compounds known as histone acetyltransferase inhibitors showed that most of them are nonselective inhibitors and either bind to thiol groups or induce enzyme aggregation. A method was proposed[228] for producing a ^{13}C-enriched reporter protein and using the ALARM-NMR approach to analyze the interaction of HTS-selected hits with the protein in order to identify selectively binding ligands.

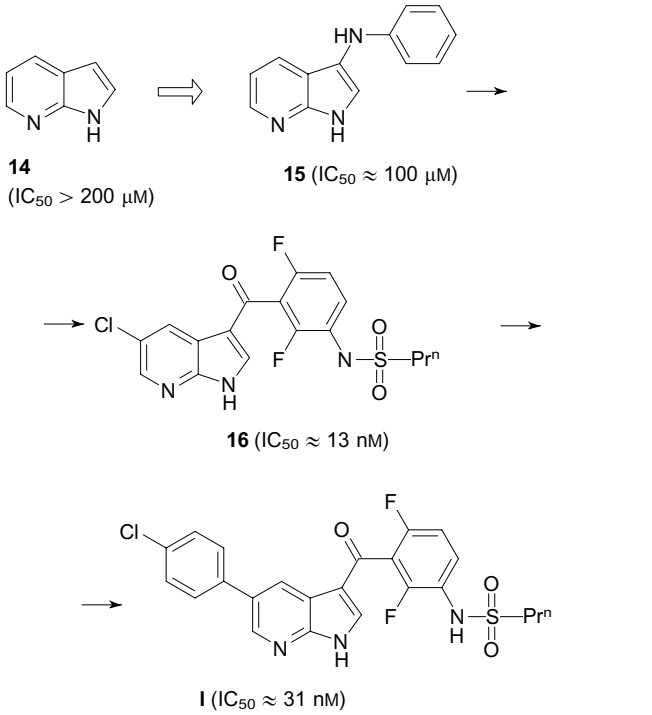
6. Examples of application of fragment-based approach and NMR methods to the design of novel drugs

Although target-oriented methods for search of new physiologically active compounds are rather young, dozens of new drugs have been designed using them.[12] Among these approaches, FBDD seems to be the youngest. Taking into account long-term obligatory clinical trials, the list of approved drugs thus designed is rather short. However, there are two approved drugs and tens of compounds that are now in late-stage clinical trials (Table 8). Besides, twenty or more compounds successfully passed preclinical testing and are now in Phase I clinical trials.[2]

6.1. Vemurafenib

Vemurafenib is a selective inhibitor of B-Raf kinases with activating mutations in the V600E codon. It was developed at Plexxikon and Genentech for the treatment of late-stage melanoma. The name 'vemurafenib' originates from V600E mutated BRAF inhibition. The V600E mutation is one of the most frequently occurring oncogenic mutations in protein kinases;[229] this is a driver mutation of the B-Raf-oncogene responsible for the development of most melanomas. A search for of selective inhibitor of V600E B-Raf kinase began at Plexxikon in the early 2005, the lead compound was found in 2011 and then approved having passed through all the phases of clinical trials. The stages of the design of vemurafenib are illustrated in Scheme 1.[230] Shown are the structures of the intermediate fragments and the lead compound (**I**) and the V600E B-Raf kinase inhibition constants (IC_{50}). The starting point, fragment **14**, has several binding sites to the target protein, while fragment **15** has only one binding site. Compound **16** binds with high affinity not only to the target, but also to other protein kinases. Eventually, compound **I** selectively binds to the V600E B-Raf kinase only. Screening of the starting fragments and their modification were carried out using biochemical methods and X-ray crystallography. The crystal structure of the complex of vemurafenib with the target, *viz.*, mutant form of V600E B-Raf kinase was established.[230]

Scheme 1



6.2 Venetoclax

Venetoclax developed at AbbVie and Genentech selectively binds to intracellular protein factors from the Bcl-2 family — regulators of programmed cell death. Unlike Vemurafenib designed within a short period of six years, the design of Venetoclax took nearly two decades. In 1996, the structure of Bcl-x_L protein, a representative member of the Bcl-2 family, was determined by NMR spectroscopy methods and X-ray crystallography.[231] Subsequent research using the SAR by NMR method over a period of a few years led to inhibitor of Bcl-x_L

Table 8. Drugs based on the lead compounds discovered using the FBLD strategy and approved or being in late-phase clinical trials.[229 – 247]

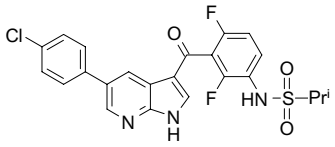
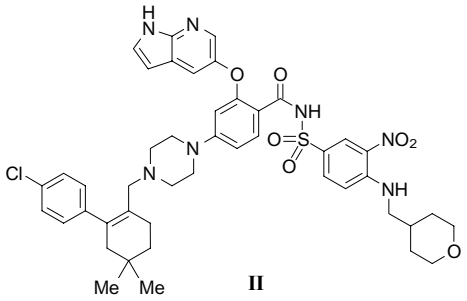
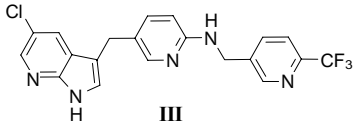
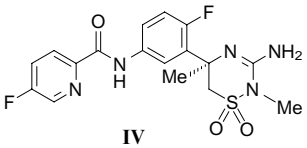
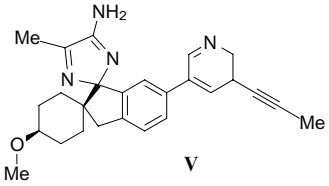
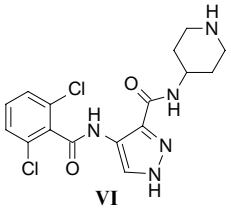
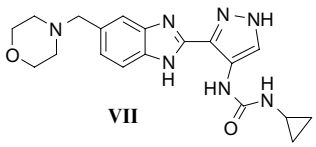
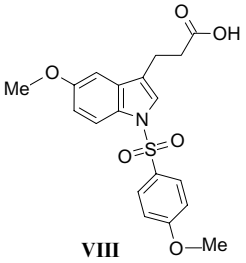
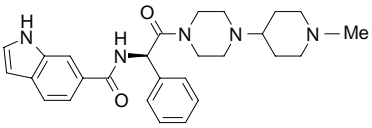
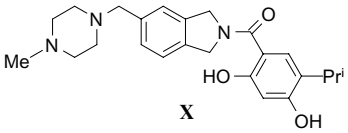
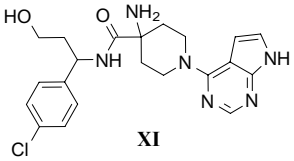
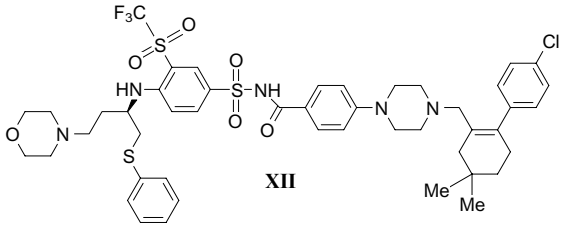
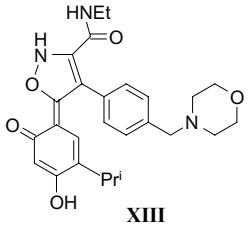
Structural formula of compound	Search method	Target	Name	Pharmaceutical company
Approved drugs				
 **I**	X-Ray crystallography	B-Raf (V600E) kinase	Vemurafenib	Plexxikon
 **II**	NMR	Bcl-2 protein — regulator of programmed cell death	Venetoclax	AbbVie, Genentech
Phase III clinical trials				
 **III**	X-Ray crystallography, NMR	Receptor of CSF1 tumour cells	Pexidartinib	Plexxikon
 **IV**	NMR	β-Secretase (BACE 1)	Verubecestat [a]	Merck
 **V**	NMR, SPR	β-Secretase (BACE 1)	Lanabecestat [a]	Astex, Lilly, AstraZeneca
Phase II clinical trials				
 **VI**	X-Ray crystallography	Cyclin-dependent kinase 2 (CDK2)	AT7519	Astex
 **VII**	X-Ray crystallography	Aurora kinase	AT9283	Astex

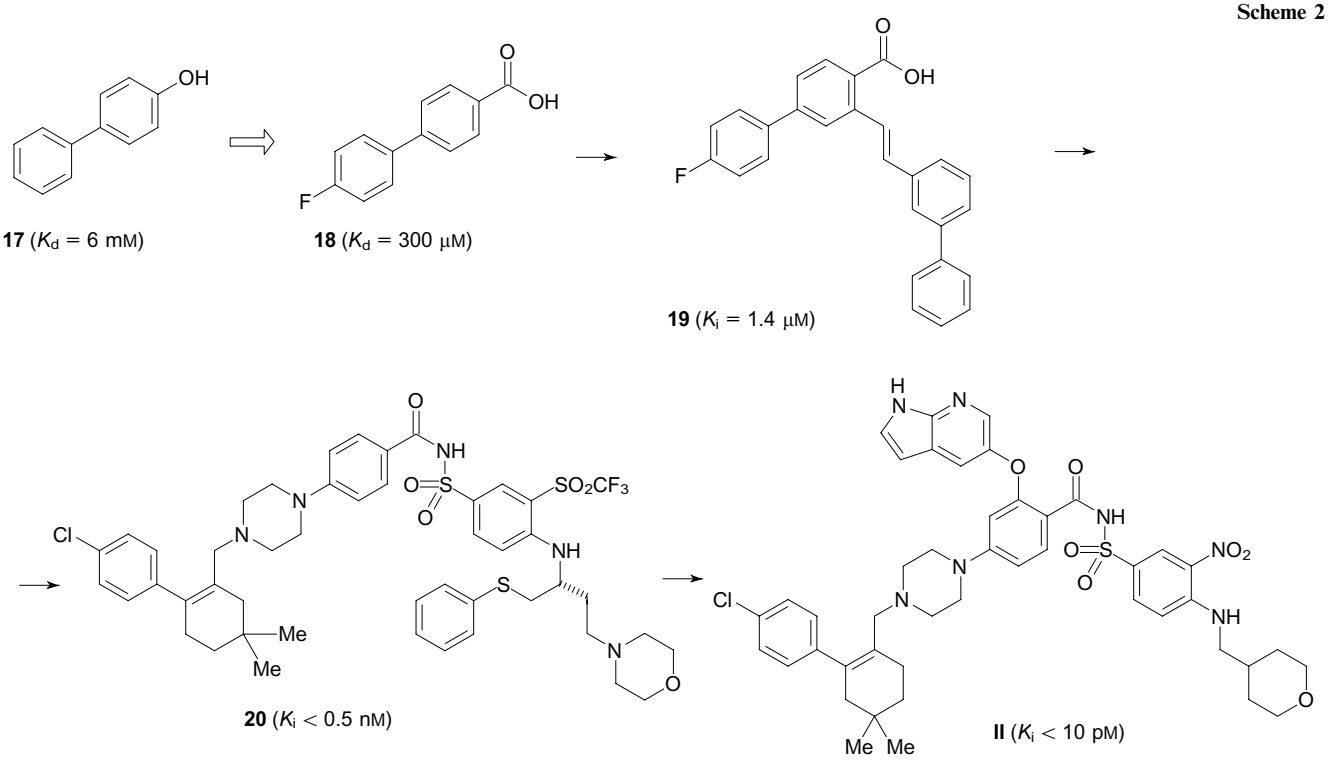
Table 8 (continued).

Structural formula of compound	Search method	Target	Name	Pharmaceutical company
Phase II clinical trials				
 **VIII**	X-Ray crystallography	Peroxisome proliferator-activated receptors Pan-PPAR	Indeglitazar	Plexxikon
 **IX**	X-Ray crystallography	Blood coagulation factor FXa	LY517717	Lilly
 **X**	NMR	Chaperone HSP90	Onalespib	Astex
 **XI**	X-Ray crystallography	Akt protein kinase	AZD5363	AstraZeneca
 **XII**	NMR	Bcl-2 protein — regulator of programmed cell death	Navitoclax	Abbott
 **XIII**	X-Ray crystallography	Chaperone HSP90	Luminespib	Vernalis

[a] For these pharmaceuticals, phase III clinical trials failed owing to low efficacy.

($K_i < 0.5$ nM, see Ref. 232). The design of Venetoclax is illustrated in Scheme 2.[232,233] The initial fragments **17** and **18** have low affinity. When linked into a single compound (structure **19**), the affinity increases by a few orders of magnitude. Further design by the NMR screening methods led to compound **20** whose treatment using medicinal chemistry methods resulted in the drug Venetoclax (**II**), a selective inhibitor of Bcl-2. The structure of compound **20** was optimized to increase the selectivity of the lead binding to Bcl-2.[233] Venetoclax (< 10 pM) inhibits Bcl-2 and is used for the treatment of chronic lymphocytic leukemia. Interestingly, Venetoclax does not obey the empirical Lipinski rules, although is administered orally.

Scheme 2



6.3. Compounds in Phase III clinical trials

By the beginning of the year 2018 three compounds designed using the FBDD approach entered Phase III clinical trials (see Table 8). These are Pexidartinib,[234] Verubecestat,[235] and Lanabecestat,[236, 237] which were designed using NMR screening methods. The starting point in the design of Pexidartinib was 7-azaindole,[238] *i.e.*, the fragment that was used by researchers at Plexxikon in the design of Vemurafenib. Pexidartinib targets CSF1 (Colony Stimulating Factor 1) receptors expressed in tumour cells. Verubecestat and Lanabecestat designed at Merck and Astex in collaboration with other pharmaceutical companies target β-secretase 1 (BACE1) involved in production of β-amyloid peptide. It is assumed that inhibition of activity of this enzyme can reduce the risk of development of Alzheimer's disease.

6.4. Compounds in Phase II clinical trials

More than ten compounds designed using the FBLD approach are in Phase II clinical trials.[2] A number of representative examples are listed in Table 8.

Compound AT7519, a cyclin-dependent kinase CDK2 inhibitor (IC_{50} = 47 nM), was developed by researchers at Astex using fragment-based X-ray crystallography.[239] It is specified as potential antitumour drug. Yet another antitumour compound designed at Astex is AT9283, an inhibitor of Aurora kinase that plays an important role in the regulation of cell proliferation.[240] Indeglitazar[241] developed at Plexxikon is a pan agonist of peroxisome proliferator-activated receptors (PPARs) — nuclear receptors involved in gene transcription processes. A possible application field of indeglitazar is the treatment of Type 2 diabetes.

Compound LY517717 was developed at Protherics Molecular Design as inhibitor of blood coagulation factor Xα involved in blood coagulation processes.[242] It is specified as potential antithrombotic agent. A potential drug onalespib (Aspex Inc.) developed using NMR screening methods is an inhibitor of heat shock protein Hsp90, a chaperone playing an important role in maintenance of tumour cell activity.[243, 244] Another Hsp90 inhibitor, luminespib (Vernalis plc) was designed using X-ray crystallography.[245]

Compound AZD5363 is a an Akt kinase (protein kinase B) inhibitor designed at AstraZeneca using fragment-based screening and X-ray crystallography.[246] The Akt kinase is a key node in the Akt signalling cascade deregulated in tumour cells. Compound AZD5363 is specified as drug for the treatment of breast cancer. Navitoclax, a structural analogue of antumour drug Venetoclax, is an inhibitor of proteins Bcl-2, Bcl-w and Bcl-x_L. It was developed by researchers at Abbot Laboratories using the NMR screening methods[247] for the treatment of lymphocytic leukemia.

7. Conclusion

Examples presented in this review illustrate the efficiency of utilizing a combination of the fragment-based strategy and NMR screening for the design of biologically active compounds. The fragment-based approach has been a commonly recognized technology of lead and drug design. A review of lead compounds thus constructed and reported before 2015 is available.[248] Among 27 compounds reviewed, eleven were designed using NMR screening methods. The fragment-based approach and NMR screening are used by researchers at many pharmaceutical companies.[249, 250] Yet another review[251] describes the application of these methods at AstraZeneca for the development of new biologically

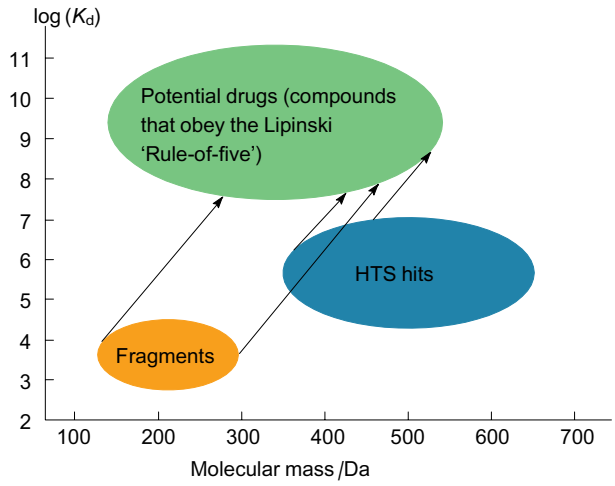


Figure 31. Assessment of approaches to the design of hit compounds based on total screening and FBLD in the coordinates K_d is molecular mass of the complex with the target.
The characteristic regions for the hits that obey the Lipinski rules, HTS hits, and molecular fragments binding to targets are given in colour. Arrow show directions of potency optimization of compounds through an increase in affinity and molecular mass. Based on the illustration from Ref. 252.

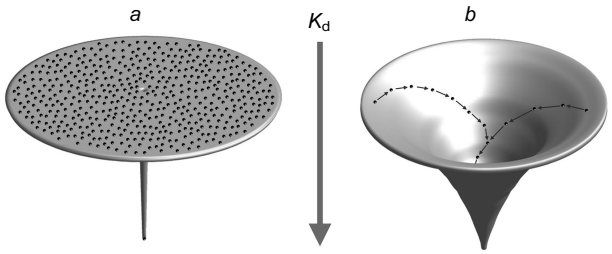


Figure 32. Schematic representation of the total screening (*a*) and fragment-based hit design (*b*) methods.
Surfaces denote the diversity of chemical structures. The depth of each point is proportional to the binding constant of corresponding compound to the target protein. Total screening aimed at finding one or more hits requires experimental studies of thousands of compounds (conditionally represented by dots). The fragment-based method combined with rational structure design strategies (including NMR screening methods) requires the synthesis and experimental studies of much smaller number of compounds. The concept of this Figure is based on the illustration from Ref. 217.

active compounds. It also covers the principles of design of fragment libraries as well as methods for selection of active fragments and subsequent lead design.

Mention should be made of differences between two modern approaches to the search and design of drugs, *viz.*, high-throughput total screening and fragment-based strategy. Figure 31 illustrates them as applied to lead design.

The plot shows regions typical of potential drugs that obey the Lipinski rules, HTS hits and structural fragments capable of binding to targets. Usually, such fragments have a mass of 150–300 Da and relatively low affinity. The affinity of the HTS hits is much higher; however, their molecular masses are also higher. To become promising lead compounds, both structural fragments and HTS hits should be modified. Structure modification of both classes of compounds is aimed at increasing their affinity to target proteins. It is accompanied by complication of the molecular structure and, most often, by an increase in the molecular mass. Thus, the optimization of compounds can be represented as the motion along the direction specified by the arrows in Fig. 31. The probability to reach the region of potential drugs is much higher for the modified fragments rather than HTS leads.

These methods of rational lead design based on NMR spectroscopy data allow one to modify molecular fragments by optimizing their properties in step-by-step manner. This feature differs the strategy in hand from the search for high-affinity compounds using HTS (Fig. 32).

High-throughput screening (see Fig. 32 *a*) requires testing thousands of compounds until a hit (or a few hits) with high affinity for the target protein are selected. In this case, the probability to reach a narrow high-affinity region is very small. The fragment-based method (Fig. 32 *b*) allows one to begin the design of a potential lead compound using one or a few fragments with relatively low affinity for the target. Subsequent fragment modification and linking will lead to an increase in the affinity of the optimized compound. By gradually varying the structure of the ligand it is possible to move along the gradient descent direction towards minimum values of the dissociation constant of the ligand–target complex. The surfaces shown in Fig. 32 denote the diversity of chemical structures while the shape of the surface is determined by the binding constants of individual compounds. Thus, the fragment-based method and NMR screening are rational approaches to the design of biologically active compounds. Undoubtedly, these methods will play an increasingly important role in the design of novel drugs.

This review has been prepared with the financial support from the Russian Science Foundation (Project No. 14-14-00598).

References

1. J.Drews. *Science*, **287**, 1960 (2000)
2. D.A.Erlanson, S.W.Fesik, R.E.Hubbard, W.Jahnke, H.Jhoti. *Nat. Rev. Drug Discov.*, **15**, 605 (2016)
3. B.Munos. *Nat. Rev. Drug Discov.*, **8**, 959 (2009)
4. H.Giersiefen, R.Hilgenfeld, A.Hillisch. *Modern Methods of Drug Discovery*. (Eds A.Hillisch, R.Hilgenfeld). (Basel, Boston, Berlin: Birkhäuser, 2003)
5. L.M.Friedman, C.Furberg, D.L.DeMets. *Fundamentals of Clinical Trials*. (New York: Springer, 2010)
6. T.L.Blundell, B.L.Sibanda, R.W.Montalvéo, S.Brewerton, V.Chelliah, C.L.Worth, N.J.Harmer, O.Davies, D.Burke. *Philos. Trans. R. Soc. London, Ser. B*, **361**, 413 (2006)

7. S.F.Kingsmore, I.E.Lindquist, J.Mudge, D.D.Gessler, W.D.Beavis. *Nat. Rev. Drug Discov.*, **7**, 221 (2008)
8. D.A.Jeffery, M.Bogyo. *Curr. Opin. Biotechnol.*, **14**, 87 (2003)
9. W.Zheng, N.Thorne, J.C.McKew. *Drug Discov. Today*, **18**, 1067 (2013)
10. J.G.Moffat, F.Vincent, J.A.Lee, J.Eder, M.Prunotto. *Nat. Rev. Drug Discov.*, **16**, 531 (2017)
11. P.Imming, C.Sinning, A.Meyer. *Nat. Rev. Drug Discov.*, **5**, 821 (2006)
12. R.Santos, O.Ursu, A.Gaulton, A.P.Bento, R.S.Donadi, C.G.Bologa, A.Karlsson, B.Al-Lazikani, A.Hersey, T.I.Oprea, J.P.Overington. *Nat. Rev. Drug Discov.*, **16**, 19 (2017)
13. D.C.Swinney, J.Anthony. *Nat. Rev. Drug Discov.*, **10**, 507 (2011)
14. J.P.Overington, B.Al-Lazikani, A.L.Hopkins. *Nat. Rev. Drug Discov.*, **5**, 993 (2006)
15. C.A.Lipinski, F.Lombardo, B.W.Dominy, P.J.Feeney. *Adv. Drug Deliv. Rev.*, **46**, 3 (2001)
16. D.F.Veber, S.R.Johnson, H.-Y.Cheng, B.R.Smith, K.W.Ward, K.D.Kopple. *J. Med. Chem.*, **45**, 2615 (2002)
17. H.van de Waterbeemd, E. Gifford. *Nat. Rev. Drug Discov.*, **2**, 192 (2003)
18. P.J.Hajduk, R.P.Meadows, S.W.Fesik. *Quart. Rev. Biophys.*, **32**, 211 (1999)
19. G.C.Roberts. *Drug Discov. Today*, **5**, 230 (2000)
20. S.S.Pochapsky, T.C.Pochapsky. *Curr. Top. Med. Chem.*, **1**, 427 (2001)
21. M.Pellecchia, D.S.Sem, K.Wuthrich. *Nat. Rev. Drug. Discov.*, **1**, 211 (2002)
22. B.J.Stockman, C.Dalvit. *Prog. Nucl. Magn. Reson. Spectrosc.*, **41**, 187 (2002)
23. B.Meyer, T.Peters. *Angew. Chem., Int. Ed.,* **42**, 864 (2003)
24. C.A.Lepre, J.M.Moore, J.W.Peng. *Chem. Rev.*, **104**, 3641 (2004)
25. J.W.Peng, J.Moore, N.Abdul-Manan. *Prog. Nucl. Magn. Reson. Spectrosc.*, **44**, 225 (2004)
26. J.Fejzo, C.Lepre, X.Xie. *Curr. Top. Med. Chem.*, **3**, 81 (2003)
27. L.Fielding. *Prog. Nucl. Magn. Reson. Spectrosc.*, **51**, 219 (2007)
28. A.Furukawa, T.Konuma, S.Yanaka, K.Sugase. *Prog. Nucl. Mag. Reson. Spectrosc.*, **96**, 47 (2016)
29. A.D.Gossert, W.Jahnke. *Prog. Nucl. Magn. Reson. Spectrosc.*, **97**, 82 (2016)
30. M.Bentley, B.C.Doak, B.Mohanty, M.J.Scanlon, in *Modern Magnetic Resonance*. (Ed G.Webb). (Basel: Springer, 2017). P. 1
31. R.Ma, P.Wang, J.Wu, K.Ruan. *Molecules*, **21**, 854 (2016)
32. D.M.Dias, A.Ciulli. *Prog. Biophys. Mol. Biol.*, **116**, 101 (2014)
33. H.Jhoti, in *Ernst Schering Foundation Symposium Proceedings. Sparking Signals. Kinases as Molecular Signaltransducers and Pharmacological Drug Targets in Inflammation*. (Berlin, Heidelberg: Springer, 2008). Vol. 3. P. 169
34. J.F.R.Mello, R.A.Gomes, D.G.V.Fujii, G.M.Ferreira, G.H.G.Trossini. *Chem. Biol. Drug Des.*, **90**, 1067 (2017)
35. G.M.Carneiro, A.B.Eiso, S.Theisgen, G.Siegal. *Essays Biochem.*, **61**, 485 (2017)
36. R.Marchetti, S.Perez, A.Arda, A.Imberty, J.Jimenez-Barbero, A.Silipo, A.Molinaro. *ChemistryOpen*, **5**, 274 (2016)
37. C.Dalvit. *Drug Discov. Today*, **14**, 1051 (2009)
38. J.P.Renaud, C.W.Chung, U.H.Danielson, U.Egner, M.Hennig, R.E.Hubbard, H.Nar. *Nat. Rev. Drug Discov.*, **15**, 679 (2016)
39. D.M.Miller, J.S.Olson, J.W.Pflugrath, F.A.Quiocho. *J. Biol. Chem.*, **258**, 3665 (1983)
40. J.Feeney, B.Birdsall, N.V.Kovalevskaya, Y.D.Smurnyy, E.M.Navarro Peran, V.I.Polshakov. *Biochemistry*, **50**, 3609 (2011)
41. A.Fersht. *Structure and Mechanism in Protein Science: a Guide to Enzyme Catalysis and Protein Folding*. (New York: W.H.Freeman and Co, 1999)
42. M.S.Marshall, R.P.Steele, K.S.Thanthiriwatte, C.D.Sherrill. *J. Phys. Chem. A*, **113**, 13628 (2009)
43. P.R.Connelly, R.A.Aldape, F.J.Bruzzese, S.P.Chambers, M.J.Fitzgibbon, M.A.Fleming, S.Itoh, D.J.Livingston, M.A.Navia, J.A.Thomson. *Proc. Natl. Acad. Sci. USA*, **91**, 1964 (1994)
44. G.B.McGaughey, M.Gagné, A.K.Rappé. *J. Biol. Chem.*, **273**, 15458 (1998)
45. D.A.Dougherty. *Science*, **271**, 163 (1996)
46. W.C.Tsui, A.R.Fersht. *Nucleic Acids Res.*, **9**, 4627 (1981)
47. M.I.Page, W.P.Jencks. *Proc. Natl. Acad. Sci. USA*, **68**, 1678 (1971)
48. A.L.Hopkins, C.R.Groom, A.Alex. *Drug Discov. Today*, **9**, 430 (2004)
49. M.L.Verdonk, D.C.Rees. *ChemMedChem*, **3**, 1179 (2008)
50. M.Rarey, in *Bioinformatics — from Genomes to Drugs*. (Ed. T.Lengauer). (Weinheim: Wiley-VCH, 2001). Vol. 1. P. 315
51. R.Macarron, M.N.Banks, D.Bojanic, D.J.Burns, D.A.Cirovic, T.Garyantes, D.V.Green, R.P.Hertzberg, W.P.Janzen, J.W.Paslay, U.Schopfer, G.S.Sittampalam. *Nat. Rev. Drug Discov.*, **10**, 188 (2011)
52. D.C.Rees, M.Congreve, C.W.Murray, R.Carr. *Nat. Rev. Drug Discov.*, **3**, 660 (2004)
53. B.J.Davis, S.D.Roughley, in *Annual Reports in Medicinal Chemistry*. (Elsevier, 2017). Vol. 50. P. 371
54. O.A.Filz, V.V.Poroikov. *Russ. Chem. Rev.*, **81**, 158 (2012)
55. A.Larsson, A.Jansson, A.Aberg, P.Nordlund. *Curr. Opin. Chem. Biol.*, **15**, 482 (2011)
56. S.M.Boyd, G.E.de Kloe. *Drug Discov. Today: Technol.*, **7**, e173 (2010)
57. M.Congreve, R.Carr, C.Murray, H.Jhoti. *Drug Discov. Today*, **8**, 876 (2003)
58. H.Jhoti, G.Williams, D.C.Rees, C.W.Murray. *Nat. Rev. Drug Discov.*, **12**, 644 (2013)
59. W.P.Jencks. *Proc. Natl. Acad. Sci. USA*, **78**, 4046 (1981)
60. S.D.Bembenek, B.A.Tounge, C.H.Reynolds. *Drug Discov. Today*, **14**, 278 (2009)
61. S.Schultes, C.de Graaf, E.E.J.Haaksma, I.J.P.de Esch, R.Leurs, O.Krämer. *Drug Discov. Today: Technol.*, **7**, e157 (2010)

62. M.J.Hartshorn, C.W.Murray, A.Cleasby, M.Frederickson, I.J.Tickle, H.Jhoti. *J. Med. Chem.*, **48**, 403 (2005)
63. L.Blundell, H.Jhoti, C.Abell. *Nat. Rev. Drug Discov.*, **1**, 45 (2002)
64. H.Jhoti, A.Cleasby, M.Verdonk, G.Williams. *Curr. Opin. Chem. Biol.*, **11**, 485 (2007)
65. M.A.Cooper. *Nat. Rev. Drug. Discov.*, **1**, 515 (2002)
66. C.A.Wartchow, F.Podlaski, S.Li, K.Rowan, X.Zhang, D.Mark, K.-S.Huang. *J. Comput. Aided Mol. Des.*, **25**, 669 (2011)
67. J.E.Ladbury, G.Klebe, E.Freire. *Nat. Rev. Drug Discov.*, **9**, 23 (2010)
68. F.H.Niesen, H.Berglund, M.Vedadi. *Nat. Protoc.*, **2**, 2212 (2007)
69. S.A.Hofstadler, K.A.Sannes-Lowery. *Nat. Rev. Drug Discov.*, **5**, 585 (2006)
70. S.Ohlson, M.-D.Duong-Thi, in *Fragment-Based Drug Discovery: Lessons and Outlook*. (Weinheim: Wiley-VCH, 2016). P. 173
71. C.J.Wienken, P.Baaske, U.Rothbauer, D.Braun, S.Duhr. *Nat. Commun.*, **1**, 100 (2010)
72. P.Baaske, C.J.Wienken, P.Reineck, S.Duhr, D.Braun. *Angew. Chem., Int. Ed.*, **49**, 2238 (2010)
73. M.P.Williamson. *Prog. Nucl. Magn. Reson. Spectrosc.*, **73**, 1 (2013)
74. M.Betz, K.Saxena, H.Schwalbe. *Curr. Opin. Chem. Biol.*, **10**, 219 (2006)
75. O.Cala, F.Guilliere, I.Krimm. *Anal. Bioanal. Chem.*, **406**, 943 (2014)
76. M.Pellecchia, I.Bertini, D.Cowburn, C.Dalvit, E.Giralt, W.Jahnke, T.L.James, S.W.Homans, H.Kessler, C.Luchinat, B.Meyer, H.Oschkinat, J.Peng, H.Schwalbe, G.Siegal. *Nat. Rev. Drug Discov.*, **7**, 738 (2008)
77. M.J.Harner, A.O.Frank, S.W.Fesik. *J. Biomol. NMR*, **56**, 65 (2013)
78. S.W.Homans. *Angew. Chem., Int. Ed.,* **43**, 290 (2004)
79. L.Banci, I.Bertini, C.Luchinat, M.Mori. *Prog. Nucl. Magn. Reson. Spectrosc.*, **56**, 247 (2010)
80. G.S.Rule, T.K.Hitchens. *Fundamentals of Protein NMR Spectroscopy. Focus on Structural Biology*. (Ed. R.Kaptein). (Dordrecht: Springer, 2006)
81. S.Y.Ohki, M.Kainosho. *Prog. Nucl. Magn. Reson. Spectrosc.*, **53**, 208 (2008)
82. V.I.Polshakov, B.Birdsall, J.Feeney. *Biochemistry*, **38**, 15962 (1999)
83. V.I.Polshakov, E.G.Smirnov, B.Birdsall, G.Kelly, J.Feeney. *J. Biomol. NMR*, **24**, 67 (2002)
84. L.Y.Lian, D.A.Middleton. *Prog. Nucl. Magn. Reson. Spectrosc.*, **39**, 171 (2001)
85. M.Ikura, A.Bax. *J. Am. Chem. Soc.*, **114**, 2433 (1992)
86. G.Gemmecker, E.T.Olejniczak, S.W.Fesik. *J. Magn. Reson.*, **96**, 199 (1992)
87. N.V.Kovalevskaya, Y.D.Smurnyy, V.I.Polshakov, B.Birdsall, A.F.Bradbury, T.Frenkiel, J.Feeney. *J. Biomol. NMR*, **33**, 69 (2005)
88. K.Michelsen, J.B.Jordan, J.Lewis, A.M.Long, E.Yang, Y.Rew, J.Zhou, P.Yakowec, P.D.Schnier, X.Huang, L.Poppe. *J. Am. Chem. Soc.*, **134**, 17059 (2012)
89. M.A.Markus, B.Dwyer, S.Wolfrom, J.C.Li, W.Li, K.Malakian, J.Wilhelm, D.H.H.Tsao. *J. Biomol. NMR*, **41**, 55 (2008)
90. V.I.Polshakov, W.D.Morgan, B.Birdsall, J.Feeney. *J. Biomol. NMR*, **14**, 115 (1999)
91. J.Orts, M.A.Wälti, M.Marsh, L.Vera, A.D.Gossert, P.Güntert, R.Riek. *J. Am. Chem. Soc.*, **138**, 4393 (2016)
92. M.Wälti, J.Orts. *Magnetochemistry*, **4**, 12 (2018)
93. V.I.Polshakov, B.Birdsall, T.A.Frenkiel, A.R.Gargaro, J.Feeney. *Protein Sci.*, **8**, 467 (1999)
94. C.J.Markin, L.Spyracopoulos. *J. Biomol. NMR*, **53**, 125 (2012)
95. R.Foster, C.A.Fyfe. *Prog. Nucl. Magn. Reson. Spectrosc.*, **4**, 1 (1969)
96. A.N.Istrate, S.A.Kozin, S.S.Zhokhov, A.B.Mantsyzov, O.I.Kechko, A.Pastore, A.A.Makarov, V.I.Polshakov. *Sci. Rep.*, **6**, 21734 (2016)
97. L.Fielding. *Curr. Top. Med. Chem.*, **3**, 39 (2003)
98. E.L.Kovrigin. *J. Biomol. NMR*, **53**, 257 (2012)
99. C.A.Waudby, A.Ramos, L.D.Cabrita, J.Christodoulou. *Sci. Rep.*, **6**, 24826 (2016)
100. U.L.Günther, B.Schaffhausen. *J. Biomol. NMR*, **22**, 201 (2002)
101. J.Krishnamoorthy, V.C.K.Yu, Y.K.Mok. *PLoS One*, **5**, e8943 (2010)
102. J.D.Hothersall, D.Guo, S.Sarda, R.J.Sheppard, H.Chen, W.Keur, M.J.Waring, A.P.I.Jzerman, S.J.Hill, I.L.Dale, P.B.Rawlins. *Mol. Pharmacol.*, **91**, 25 (2017)
103. C.J.Craven, B.Whitehead, S.K.A.Jones, E.Thulin, G.M.Blackburn, J.P.Waltho. *Biochemistry*, **35**, 10287 (1996)
104. P.E.Johnson, A.L.Creagh, E.Brun, K.Joe, P.Tomme, C.A.Haynes, L.P.McIntosh. *Biochemistry*, **37**, 12772 (1998)
105. F.A.A.Mulder, N.R.Skrynnikov, B.Hon, F.W.Dahlquist, L.E.Kay. *J. Am. Chem. Soc.*, **123**, 967 (2001)
106. M.Tollinger, N.R.Skrynnikov, F.A.A.Mulder, J.D.Forman-Kay, L.E.Kay. *J. Am. Chem. Soc.*, **123**, 11341 (2001)
107. E.Walinda, D.Morimoto, K.Sugase. *Curr. Protoc. Prot. Sci.*, **92**, e57 (2018)
108. D.Tolkatchev, P.Xu, F.Ni. *J. Am. Chem. Soc.*, **125**, 12432 (2003)
109. T.Moschen, S.Grutsch, M.A.Juen, C.H.Wunderlich, C.Kreutz, M.Tollinger. *J. Med. Chem.*, **59**, 10788 (2016)
110. T.Moschen, C.H.Wunderlich, R.Spitzer, J.Levic, R.Micura, M.Tollinger, C.Kreutz. *Angew. Chem., Int. Ed.*, **127**, 570 (2015)
111. J.S.Zintsmaster, B.D.Wilson, J.W.Peng. *J. Am. Chem. Soc.*, **130**, 14060 (2008)
112. P.-C.Lin. *J. Biomol. NMR*, **62**, 121 (2015)
113. P.Trigo-Mouriåo, C.Griesinger, D.Lee. *J. Biomol. NMR*, **69**, 229 (2017)
114. P.Job. *Ann. Chim. Fr.*, **9**, 113 (1928)
115. J.S.Renny, L.L.Tomasevich, E.H.Tallmadge, D.B.Collum. *Angew. Chem., Int. Ed.*, **52**, 11998 (2013)
116. K.Klausen. *Anal. Chim. Acta*, **44**, 377 (1969)
117. V.M.S.Gil, N.C.Oliveira. *J. Chem. Educ.*, **67**, 473 (1990)

118. S.A.Kozin, A.A.Kulikova, A.N.Istrate, P.O.Tsvetkov, S.S.Zhokhov, Y.V.Mezentsev, O.I.Kechko, A.S.Ivanov, V.I.Polshakov, A.A.Makarov. *Metallomics*, **7**, 422 (2015)
119. P.Luginbühl, K.Wüthrich. *Prog. Nucl. Magn. Reson. Spectrosc.*, **40**, 199 (2002)
120. P.J.Hajduk, E.T.Olejniczak, S.W.Fesik. *J. Am. Chem. Soc.*, **119**, 12257 (1997)
121. S.V.Kharlamov, S.K.Latypov. *Russ. Chem. Rev.*, **79**, 635 (2010)
122. M.Lin, M.J.Shapiro, J.R.Wareing. *J. Am. Chem. Soc.*, **119**, 5249 (1997)
123. S.B.Shuker, P.J.Hajduk, R.P.Meadows, S.W.Fesik. *Science*, **274**, 1531 (1996)
124. P.J.Hajduk, G.Sheppard, D.G.Nettesheim, E.T.Olejniczak, S.B.Shuker, R.P.Meadows, D.H.Steinman, G.M.Carrera, P.A.Marcotte, J.Severin, K.Walter, H.Smith, E.Gubbins, R.Simmer, T.F.Holzman, D.W.Morgan, S.K.Davidsen, J.B.Summers, S.W.Fesik. *J. Am. Chem. Soc.*, **119**, 5818 (1997)
125. P.Schanda, E.Kupče, B.Brutscher. *J. Biomol. NMR*, **33**, 199 (2005)
126. K.Pervushin, R.Riek, G.Wider, K.Wüthrich. *Proc. Natl. Acad. Sci. USA*, **94**, 12366 (1997)
127. R.Riek, K.Pervushin, K.Wüthrich. *Tr. Biochem. Sci.*, **25**, 462 (2000)
128. P.J.Hajduk, D.J.Augeri, J.Mack, R.Mendoza, J.Yang, S.F.Betz, S.W.Fesik. *J. Am. Chem. Soc.*, **122**, 7898 (2000)
129. A.M.Ruschak, L.E.Kay. *J. Biomol. NMR*, **46**, 75 (2009)
130. A.M.Ruschak, M.Slassi, L.E.Kay, A.D.Schimmer. *JNCI: J. Nat. Cancer Inst.*, **103**, 1007 (2011)
131. P.J.Hajduk. *Mol. Interv.*, **6**, 266 (2006)
132. E.R.Zartler, J.Hanson, B.E.Jones, A.D.Kline, G.Martin, H.Mo, M.J.Shapiro, R.Wang, H.Wu, J.Yan. *J. Am. Chem. Soc.*, **125**, 10941 (2003)
133. D.A.Vinarov, B.L.Lytle, F.C.Peterson, E.M.Tyler, B.F.Volkman, J.L.Markley. *Nat. Methods*, **1** (2), 149 (2004)
134. M.J.Parker, M.Aulton-Jones, A.M.Hounslow, C.J.Craven. *J. Am. Chem. Soc.*, **126**, 5020 (2004)
135. N.K.Mishra, A.K.Urick, S.W.J.Ember, E.Schönbrunn, W.C.Pomerantz. *ACS Chem. Biol.*, **9**, 2755 (2014)
136. C.T.Gee, K.E.Arntson, A.K.Urick, N.K.Mishra, L.M.L.Hawk, A.J.Wisniewski, W.C.K.Pomerantz. *Nat. Protocols*, **11**, 1414 (2016)
137. A.K.Urick, L.P.Calle Jimenez, J.F.Espinosa, H.Hu, W.C.Pomerantz. *ACS Chem. Biol.*, **11**, 3154 (2016)
138. C.T.Gee, E.J.Koleski, W.C.K.Pomerantz. *Angew. Chem., Int. Ed.*, **54**, 3735 (2015)
139. E.W.W.Leung, H.Yagi, J.R.Harjani, M.D.Mulcair, M.J.Scanlon, J.B.Baell, R.S.Norton. *Chem. Biol. Drug. Des.*, **84**, 616 (2014)
140. X.Ge, C.A.MacRaild, S.M.Devine, C.O.Debono, G.Wang, P.J.Scammells, M.J.Scanlon, R.F.Anders, M.Foley, R.S.Norton. *J. Med. Chem.*, **57**, 6419 (2014)
141. W.Jahnke, L.B.Perez, C.G.Paris, A.Strauss, G.Fendrich, C.M.Nalin. *J. Am. Chem. Soc.*, **122**, 7394 (2000)
142. G.Otting. *Annu. Rev. Biophys.*, **39**, 387 (2010)
143. U.Brath, S.I.Swamy, A.X.Veiga, C.-C.Tung, F.Van Petegem, M.Erdélyi. *J. Am. Chem. Soc.*, **137**, 11391 (2015)
144. G.Zhu, X.J.Yao. *Prog. Nucl. Magn. Reson. Spectrosc.*, **52**, 49 (2008)
145. C.Caillet-Saguy, M.Piccioli, P.Turano, N.Izadi-Pruneyre, M.Delepierre, I.Bertini, A.Lecroisey. *J. Am. Chem. Soc.*, **131**, 1736 (2009)
146. M.Mayer, B.Meyer. *Angew. Chem., Int. Ed.*, **38**, 1784 (1999)
147. B.J.Davis, A.M.Giannetti, in *Fragment-Based Drug Discovery Lessons and Outlook*. (Weinheim: Wiley-VCH, 2016). P. 119
148. H.-Y.Kim, D.F.Wyss, in *Chemical Biology: Methods and Protocols*. (Eds E.J.Hempel, H. C.Williams, C.C.Hong). (New York: Springer, 2015). P. 197
149. M.Webba da Silva. *Methods*, **43**, 264 (2007)
150. B.Cutting, S.V.Shelke, Z.Dragic, B.Wagner, H.Gathje, S.Kelm, B.Ernst. *Magn. Reson. Chem.*, **45**, 720 (2007)
151. J.L.Wagstaff, S.Vallath, J.F.Marshall, R.A.Williamson, M.J.Howard. *Chem. Commun.*, **46**, 7533 (2010)
152. J.L.Wagstaff, S.L.Taylor, M.J.Howard. *Mol. Biosyst.*, **9**, 571 (2013)
153. J.Angulo, P.M.Enríquez-Navas, P.M. Nieto. *Chem. – Eur. J.*, **16**, 7803 (2010)
154. J.Angulo, P.M.Nieto. *Eur. Biophys. J.*, **40**, 1357 (2011)
155. A.Viegas, J.Manso, F.L.Nobrega, E.J.Cabrita. *J. Chem. Educ.*, **88**, 990 (2011)
156. T.Diercks, J.P.Ribeiro, F.J.Caåada, S.André, J.Jiménez-Barbero, H.-J.Gabius. *Chem. – Eur. J.*, **15**, 5666 (2009)
157. B.Claasen, M.Axmann, R.Meinecke, B.Meyer. *J. Am. Chem. Soc.*, **127**, 916 (2005)
158. T.Haselhorst, A.K.Münster-Kühnel, M.Oschlies, J.Tiralongo, R.Gerardy-Schahn, M.v.Itzstein. *Biochem. Biophys. Res. Commun.*, **359**, 866 (2007)
159. C.Dalvit, G.Fogliatto, A.Stewart, M.Veronesi, B.Stockman. *J. Biomol. NMR*, **21**, 349 (2001)
160. A.D.Gossert, C.Henry, M.J.Blommers, W.Jahnke, C.Fernandez. *J. Biomol. NMR*, **43**, 211 (2009)
161. J.Hu, P.O.Eriksson, G.Kern. *Magn. Reson. Chem.*, **48**, 909 (2010)
162. A.Antanasijevic, B.Ramirez, M.Caffrey. *J. Biomol. NMR*, **60**, 37 (2014)
163. K.Furihata, H.Utsumi, T.Kato, C.Sakuma, M.Tashiro. *Pharm. Anal. Chem. (Open Access)*, **2**, 111 (2016)
164. E.A.Batuev, A.Y.Lisunov, E.A.Morozova, V.V.Klochkov, N.V.Anufrieva, T.V.Demidkina, V.I.Polshakov. *Mol. Biol.*, **48**, 896 (2014)
165. L.Fielding, S.Rutherford, D.Fletcher. *Magn. Reson. Chem.*, **43**, 463 (2005)
166. R.Huang, A.Bonnichon, T.D.W.Claridge, I.K.H.Leung. *Sci. Rep.*, **7**, 43727 (2017)
167. C.Dalvit, P.E.Fagerness, D.T.Hadden, R.W.Sarver, B.J.Stockman. *J. Am. Chem. Soc.*, **125**, 7696 (2003)
168. C.Dalvit, N.Mongelli, G.Papeo, P.Giordano, M.Veronesi, D.Moskau, R.Kummerle. *J. Am. Chem. Soc.*, **127**, 13380 (2005)
169. C.Dalvit, E.Ardini, G.P.Fogliatto, N.Mongelli, M.Veronesi. *Drug Discov. Today*, **9**, 595 (2004)

170. C.Dalvit. *Prog. Nucl. Magn. Reson. Spectrosc.*, **51**, 243 (2007)
171. C.Lambruschini, M.Veronesi, E.Romeo, G.Garau, T.Bandiera, D.Piomelli, R.Scarpelli, C.Dalvit. *ChemBioChem*, **14**, 1611 (2013)
172. M.Veronesi, F.Giacomina, E.Romeo, B.Castellani, G.Ottonello, C.Lambruschini, G.Garau, R.Scarpelli, T.Bandiera, D.Piomelli, C.Dalvit. *Anal. Biochem.*, **495**, 52 (2016)
173. W.Jahnke, S.Rudisser, M.Zurini. *J. Am. Chem. Soc.*, **123**, 3149 (2001)
174. W.Jahnke. *ChemBioChem*, **3**, 167 (2002)
175. B.Liang, J.H.Bushweller, L.K.Tamm. *J. Am. Chem. Soc.*, **128**, 4389 (2006)
176. C.Tang, C.D.Schwieters, G.M.Clore. *Nature (London)*, **449**, 1078 (2007)
177. A.Ramos, G.Varani. *J. Am. Chem. Soc.*, **120**, 10992 (1998)
178. J.Iwahara, C.D.Schwieters, G.M.Clore. *J. Am. Chem. Soc.*, **126**, 5879 (2004)
179. S.Vanwetswinkel, R.J.Heetebrij, J.van Duynhoven, J.G.Hollander, D.V.Filippov, P.J.Hajduk, G.Siegal. *Chem. Biol.*, **12**, 207 (2005)
180. N.M.Loening, J.Keeler, G.A.Morris. *J. Magn. Reson.*, **153**, 103 (2001)
181. M.D.Pelta, H.Barjat, G.A.Morris, A.L.Davis, S.J.Hammond. *Magn. Reson. Chem.*, **36**, 706 (1998)
182. A.Chen, D.Wu, C.S.Johnson Jr. *J. Phys. Chem.*, **99**, 828 (1995)
183. V.M.Sanchez-Pedregal, M.Reese, J.Meiler, M.J.Blommers, C.Griesinger, T.Carlomagno. *Angew. Chem., Int. Ed.*, **44**, 4172 (2005)
184. B.Becattini, M.Pellecchia. *Chemistry*, **12**, 2658 (2006)
185. B.Becattini, S.Sareth, D.Zhai, K.J.Crowell, M.Leone, J.C.Reed, M.Pellecchia. *Chem. Biol.*, **11**, 1107 (2004)
186. C.Ludwig, P.J.Michiels, X.Wu, K.L.Kavanagh, E.Pilka, A.Jansson, U.Oppermann, U.L.Gunther. *J. Med. Chem.*, **51** (1), 1 (2008)
187. M.Mayer, B.Meyer. *J. Am. Chem. Soc.*, **123**, 6108 (2001)
188. J.Orts, C.Griesinger, T.Carlomagno. *J. Magn. Reson.*, **200**, 64 (2009)
189. I.Gómez Pinto, C.Guilbert, N.B.Ulyanov, J.Stearns, T.L.James. *J. Med. Chem.*, **51**, 7205 (2008)
190. L.Skjærven, L.Codutti, A.Angelini, M.Grimaldi, D.Latek, P.Monecke, M.K.Dreyer, T.Carlomagno. *J. Am. Chem. Soc.*, **135**, 5819 (2013)
191. M.Reese, V.M.Sanchez-Pedregal, K.Kubicek, J.Meiler, M.J.J.Blommers, C.Griesinger, T.Carlomagno. *Angew. Chem., Int. Ed. Engl.*, **46**, 1864 (2007)
192. K.Ono, K.Takeuchi, H.Ueda, Y.Morita, R.Tanimura, I.Shimada, H.Takahashi. *Angew. Chem., Int. Ed.*, **53**, 2597 (2014)
193. A.Sivertsen, J.Isaksson, H.-K.Leiros, J.Svenson, J.-S.Svendsen, B.Brandsdal. *BMC Struct. Biol.*, **14**, 4 (2014)
194. I.Krimm. *Med. Chem. Commun.*, 605 (2012)
195. M.Fruth, A.Plaza, S.Hinsberger, J.H.Sahner, J.Haupenthal, M.Bischoff, R.Jansen, R.Müller, R.W.Hartmann. *ACS Chem. Biol.*, **9**, 2656 (2014)
196. M.Pellecchia, B.Becattini, K.J.Crowell, R.Fattorusso, M.Forino, M.Fragai, D.Jung, T.Mustelin, L.Tautz. *Exp. Opin. Ther. Tar.*, **8**, 597 (2004)
197. D.Li, E.F.DeRose, R.E.London. *J. Biomol. NMR*, **15**, 71 (1999)
198. C.Aguirre, O.Cala, I.Krimm. *Curr. Protoc. Protein Sci.*, **81** (1), 17.18 (2015)
199. B.Becattini, C.Culmsee, M.Leone, D.Zhai, X.Zhang, K.J.Crowell, M.F.Rega, S.Landshamer, J.C.Reed, N.Plesnila, M.Pellecchia. *Proc. Natl. Acad. Sci. USA*, **103**, 12602 (2006)
200. J.Yan, A.D.Kline, H.Mo, M.J.Shapiro, E.R.Zartler. *J. Magn. Res.*, **163**, 270 (2003)
201. V.Jayalakshmi, N.R.Krishna. *J. Magn. Reson.*, **155**, 106 (2002)
202. W.Zhang, R.Li, R.Shin, Y.Wang, I.Padmalayam, L.Zhai, N.R.Krishna. *ChemMedChem*, **8**, 1629 (2013)
203. Y.Yuan, X.Wen, D.A.R.Sanders, B.M.Pinto. *Biochemistry*, **44**, 14080 (2005)
204. J.A.Kelly, J.Feigon, T.O.Yeates. *J. Mol. Biol.*, **256**, 417 (1996)
205. P.Schultze, R.F.Macaya, J.Feigon. *J. Mol. Biol.*, **235**, 1532 (1994)
206. K.Y.Wang, S.H.Krawczyk, N.Bischofberger, S.Swaminathan, P.H.Bolton. *Biochemistry*, **32**, 11285 (1993)
207. R.F.Macaya, P.Schultze, F.W.Smith, J.A.Roe, J.Feigon. *Proc. Natl. Acad. Sci. USA*, **90**, 3745 (1993)
208. S.Uesugi, T.Kaneyasu, J.Imura, M.Ikehara, D.M.Cheng, L.S.Kan, P.O.P.Tso. *Biopolymers*, **22**, 1189 (1983)
209. C.Peikert, K.Seeger, R.K.Bhat, S.Berger. *Org. Biomol. Chem.*, **2**, 1777 (2004)
210. P.J.Hajduk, J.C.Mack, E.T.Olejniczak, C.Park, P.J.Dandliker, B.A.Beutel. *J. Am. Chem. Soc.*, **126**, 2390 (2004)
211. R.E.London. *J. Magn. Reson.*, **141**, 301 (1999)
212. H.-C.Siebert, J.Jiménez-Barbero, S.André, H.Kaltner, H.-J.Gabius, in *Methods in Enzymology. Recognition of Carbohydrates in Biological Systems. Part A: General Procedures*. (Elsevier, 2003). Vol. 362. P. 417
213. G.M.Clore, A.M.Gronenborn. *J. Magn. Reson.*, **53**, 423 (1983)
214. H.Furukawa, T.Hamada, M.K.Hayashi, T.Haga, Y.Muto, H.Hirota, S.Yokoyama, K.Nagasawa, M.Ishiguro. *Mol. Pharmacol.*, **62**, 778 (2002)
215. M.Mayer, B.Meyer. *J. Med. Chem.*, **43**, 2093 (2000)
216. B.Meyer, T.Weimar, T.Peters. *Eur. J. Biochem.*, **246**, 705 (1997)
217. J.Fejzo, C.A.Lepre, J.W.Peng, G.W.Bemis, Ajay, M.A.Murcko, J.M.Moore. *Chem. Biol.*, **6**, 755 (1999)
218. J.W.Peng, C.A.Lepre, J.Fejzo, N.Abdul-Manan, J.M.Moore, in *Methods in Enzymology. Nuclear Magnetic Resonance of Biological Macromolecules. Part A*. (Academic Press, 2001). Vol. 338. P. 202
219. E.C.Johnson, V.A.Feher, J.W.Peng, J.M.Moore, J.R.Williamson. *J. Am. Chem. Soc.*, **125**, 15724 (2003)
220. X.Arroyo, M.Goldflam, M.Feliz, I.Belda, E.Giralt. *PLoS One*, **8**, e58571 (2013)

221. J.L.Stark, H.R.Eghbalnia, W.Lee, W.M.Westler, J.L.Markley. *J. Proteome Res.*, **15**, 1360 (2016)
222. P.J.Hajduk, M.Bures, J.Praestgaard, S.W.Fesik. *J. Med. Chem.*, **43**, 3443 (2000)
223. M.Pellecchia, D.Meininger, Q.Dong, E.Chang, R.Jack, D.S.Sem. *J. Biomol. NMR*, **22**, 165 (2002)
224. G.M.Rishton. *Drug Discov. Today*, **8**, 86 (2003)
225. J.R.Huth, R.Mendoza, E.T.Olejniczak, R.W.Johnson, D.A.Cothron, Y.Liu, C.G.Lerner, J.Chen, P.J.Hajduk. *J. Am. Chem. Soc.*, **127**, 217 (2005)
226. E.R.Zartler, M.J.Shapiro. *Curr. Opin. Chem. Biol.*, **9**, 366 (2005)
227. J.L.Dahlin, K.M.Nelson, J.M.Strasser, D.Barsyte-Lovejoy, M.M.Szewczyk, S.Organ, M.Cuellar, G.Singh, J.H.Shrimp, N.Nguyen, J.L.Meier, C.H.Arrowsmith, P.J.Brown, J.B.Baell, M.A.Walters. *Nat. Commun.*, **8**, 1527 (2017)
228. J.L.Dahlin, M.Cuellar, G.Singh, K.M.Nelson, J.M.Strasser, T.Rappe, Y.Xia, G.Veglia, M.A.Walters. *Curr. Prot. Chem. Biol.*, **10**, 91 (2018)
229. J.Tsai, J.T.Lee, W.Wang, J.Zhang, H.Cho, S.Mamo, R.Bremer, S.Gillette, J.Kong, N.K.Haass, K.Sproesser, L.Li, K.S.M.Smalley, D.Fong, Y.-L.Zhu, A.Marimuthu, H.Nguyen, B.Lam, J.Liu, I.Cheung, J.Rice, Y.Suzuki, C.Luu, C.Settachatgul, R.Shellooe, J.Cantwell, S.-H.Kim, J.Schlessinger, K.Y.J.Zhang, B.L.West, B.Powell, G.Habets, C.Zhang, P.N.Ibrahim, P.Hirth, D.R.Artis, M.Herlyn, G.Bollag. *Proc. Natl. Acad. Sci. USA*, **105**, 3041 (2008)
230. G.Bollag, P.Hirth, J.Tsai, J.Zhang, P.N.Ibrahim, H.Cho, W.Spevak, C.Zhang, Y.Zhang, G.Habets, E.A.Burton, B.Wong, G.Tsang, B.L.West, B.Powell, R.Shellooe, A.Marimuthu, H.Nguyen, K.Y.J.Zhang, D.R.Artis, J.Schlessinger, F.Su, B.Higgins, R.Iyer, K.D'Andrea, A.Koehler, M.Stumm, P.S.Lin, R.J.Lee, J.Grippo, I.Puzanov, K.B.Kim, A.Ribas, G.A.McArthur, J.A.Sosman, P.B.Chapman, K.T.Flaherty, X.Xu, K.L.Nathanson, K.Nolop. *Nature (London)*, **467**, 596 (2010)
231. S.W.Muchmore, M.Sattler, H.Liang, R.P.Meadows, J.E.Harlan, H.S.Yoon, D.Nettesheim, B.S.Chang, C.B.Thompson, S.-L.Wong, S.-C.Ng, S.W.Fesik. *Nature (London)*, **381**, 335 (1996)
232. T.Oltersdorf, S.W.Elmore, A.R.Shoemaker, R.C.Armstrong, D.J.Augeri, B.A.Belli, M.Bruncko, T.L.Deckwerth, J.Dinges, P.J.Hajduk, M.K.Joseph, S.Kitada, S.J.Korsmeyer, A.R.Kunzer, A.Letai, C.Li, M.J.Mitten, D.G.Nettesheim, S.Ng, P.M.Nimmer, J.M.O'Connor, A.Oleksijew, A.M.Petros, J.C.Reed, W.Shen, S.K.Tahir, C.B.Thompson, K.J.Tomaselli, B.Wang, M.D.Wendt, H.Zhang, S.W.Fesik, S.H.Rosenberg. *Nature (London)*, **435**, 677 (2005)
233. A.J.Souers, J.D.Leverson, E.R.Boghaert, S.L.Ackler, N.D.Catron, J.Chen, B.D.Dayton, H.Ding, S.H.Enschede, W.J.Fairbrother, D.C.S.Huang, S.G.Hymowitz, S.Jin, S.L.Khaw, P.J.Kovar, L.T.Lam, J.Lee, H.L.Maecker, K.C.Marsh, K.D.Mason, M.J.Mitten, P.M.Nimmer, A.Oleksijew, C.H.Park, C.-M.Park, D.C.Phillips, A.W.Roberts, D.Sampath, J.F.Seymour, M.L.Smith, G.M.Sullivan, S.K.Tahir, C.Tse, M.D.Wendt, Y.Xiao, J.C.Xue, H.Zhang, R.A.Humerickhouse, S.H.Rosenberg, S.W.Elmore. *Nat. Med.*, **19**, 202 (2013)
234. W.D.Tap, Z.A.Wainberg, S.P.Anthony, P.N.Ibrahim, C.Zhang, J.H.Healey, B.Chmielowski, A.P.Staddon, A.L.Cohn, G.I.Shapiro, V.L.Keedy, A.S.Singh, I.Puzanov, E.L.Kwak, A.J.Wagner, D.D.Von Hoff, G.J.Weiss, R.K.Ramanathan, J.Zhang, G.Habets, Y.Zhang, E.A.Burton, G.Visor, L.Sanftner, P.Severson, H.Nguyen, M.J.Kim, A.Marimuthu, G.Tsang, R.Shellooe, C.Gee, B.L.West, P.Hirth, K.Nolop, M.van de Rijn, H.H.Hsu, C.Peterfy, P.S.Lin, S.Tong-Starksen, G.Bollag. *New Engl. J. Med.*, **373**, 428 (2015)
235. J.D.Scott, S.W.Li, A.P.J.Brunskill, X.Chen, K.Cox, J.N.Cumming, M.Forman, E.J.Gilbert, R.A.Hodgson, L.A.Hyde, Q.Jiang, U.Iserloh, I.Kazakevich, R.Kuvelkar, H.Mei, J.Meredith, J.Misiaszek, P.Orth, L.M.Rossiter, M.Slater, J.Stone, C.O.Strickland, J.H.Voigt, G.Wang, H.Wang, Y.Wu, W.J.Greenlee, E.M.Parker, M.E.Kennedy, A.W.Stamford. *J. Med. Chem.*, **59**, 10435 (2016)
236. D.Oehlrich, H.Prokopcova, H.J.M.Gijsen. *Bioorg. Med. Chem. Lett.*, **24**, 2033 (2014)
237. S.Eketjäll, J.Janson, K.Kaspersson, A.Bogstedt, F.Jeppsson, J.Fälting, S.B.Haeberlein, A.R.Kugler, R.C.Alexander, G.Cebers. *J. Alzheimers Dis.*, **50**, 1109 (2016)
238. C.Zhang, P.N.Ibrahim, J.Zhang, E.A.Burton, G.Habets, Y.Zhang, B.Powell, B.L.West, B.Matusow, G.Tsang, R.Shellooe, H.Carias, H.Nguyen, A.Marimuthu, K.Y.J.Zhang, A.Oh, R.Bremer, C.R.Hurt, D.R.Artis, G.Wu, M.Nespi, W.Spevak, P.Lin, K.Nolop, P.Hirth, G.H.Tesch, G.Bollag. *Proc. Natl. Acad. Sci. USA*, **110**, 5689 (2013)
239. P.G.Wyatt, A.J.Woodhead, V.Berdini, J.A.Boulstridge, M.G.Carr, D.M.Cross, D.J.Davis, L.A.Devine, T.R.Early, R.E.Feltell, E.J.Lewis, R.L.McMenamin, E.F.Navarro, M.A.O'Brien, M.O'Reilly, M.Reule, G.Saxty, L.C.A.Seavers, D.-M.Smith, M.S.Squires, G.Trewartha, M.T.Walker, A.J.A.Woolford. *J. Med. Chem.*, **51**, 4986 (2008)
240. S.Howard, V.Berdini, J.A.Boulstridge, M.G.Carr, D.M.Cross, J.Curry, L.A.Devine, T.R.Early, L.Fazal, A.L.Gill, M.Heathcote, S.Maman, J.E.Matthews, R.L.McMenamin, E.F.Navarro, M.A.O'Brien, M.O'Reilly, D.C.Rees, M.Reule, D.Tisi, G.Williams, M.Vinković, P.G.Wyatt. *J. Med. Chem.*, **52**, 379 (2009)
241. D.R.Artis, J.J.Lin, C.Zhang, W.Wang, U.Mehra, M.Perreault, D.Erbe, H.I.Krupka, B.P.England, J.Arnold, A.N.Plotnikov, A.Marimuthu, H.Nguyen, S.Will, M.Signaevsky, J.Kral, J.Cantwell, C.Settachatgull, D.S.Yan, D.Fong, A.Oh, S.Shi, P.Womack, B.Powell, G.Habets, B.L.West, K.Y.J.Zhang, M.V.Milburn, G.P.Vlasuk, K.P.Hirth, K.Nolop, G.Bollag, P.N.Ibrahim, J.F.Tobin. *Proc. Natl. Acad. Sci. USA*, **106**, 262 (2009)

242. J.W.Liebeschuetz, S.D.Jones, P.J.Morgan, C.W.Murray, A.D.Rimmer, J.M.E.Roscoe, B.Waszkowycz, P.M.Welsh, W.A.Wylie, S.C.Young, H.Martin, J.Mahler, L.Brady, K.Wilkinson. *J. Med. Chem.*, **45**, 1221 (2002)
243. A.J.Woodhead, H.Angove, M.G.Carr, G.Chessari, M.Congreve, J.E.Coyle, J.Cosme, B.Graham, P.J.Day, R.Downham, L.Fazal, R.Feltell, E.Figuero, M.Frederickson, J.Lewis, R.McMenamin, C.W.Murray, M.A.O'Brien, L.Parra, S.Patel, T.Phillips, D.C.Rees, S.Rich, D.-M.Smith, G.Trewartha, M.Vinkovic, B.Williams, A.J.A.Woolford. *J. Med. Chem.*, **53**, 5956 (2010)
244. C.W.Murray, M.G.Carr, O.Callaghan, G.Chessari, M.Congreve, S.Cowan, J.E.Coyle, R.Downham, E.Figueroa, M.Frederickson, B.Graham, R.McMenamin, M.A.O'Brien, S.Patel, T.R.Phillips, G.Williams, A.J.Woodhead, A.J.A.Woolford. *J. Med. Chem.*, **53**, 5942 (2010)
245. P.A.Brough, W.Aherne, X.Barril, J.Borgognoni, K.Boxall, J.E.Cansfield, K.-M.J.Cheung, I.Collins, N.G.M.Davies, M.J.Drysdale, B.Dymock, S.A.Eccles, H.Finch, A.Fink, A.Hayes, R.Howes, R.E.Hubbard, K.James, A.M.Jordan, A.Lockie, V.Martins, A.Massey, T.P.Matthews, E.McDonald, C.J.Northfield, L.H.Pearl, C.Prodromou, S.Ray, F.I.Raynaud, S.D.Roughley, S.Y.Sharp, A.Surgenor, D.L.Walmsley, P.Webb, M.Wood, P.Workman, L.Wright. *J. Med. Chem.*, **51**, 196 (2008)
246. M.Addie, P.Ballard, D.Buttar, C.Crafter, G.Currie, B.R.Davies, J.Debreczeni, H.Dry, P.Dudley, R.Greenwood, P.D.Johnson, J.G.Kettle, C.Lane, G.Lamont, A.Leach, R.W.A.Luke, J.Morris, D.Ogilvie, K.Page, M.Pass, S.Pearson, L.Ruston. *J. Med. Chem.*, **56**, 2059 (2013)
247. M.D.Wendt, in *Protein–Protein Interactions*. (Ed. M.D.Wendt). (Berlin, Heidelberg: Springer, 2012). P. 231
248. C.N.Johnson, D.A.Erlanson, C.W.Murray, D.C.Rees. *J. Med. Chem.*, **60**, 89 (2017)
249. R.Campos-Olivas. *Curr. Top. Med. Chem.*, **11**, 43 (2011)
250. G.M.Keseru, G.M.Makara. *Drug Discov. Today*, **11**, 741 (2006)
251. N.Fuller, L.Spadola, S.Cowen, J.Patel, H.Schönherr, Q.Cao, A.McKenzie, F.Edfeldt, A.Rabow, R.Goodnow. *Drug Discov. Today*, **21**, 1272 (2016)
252. R.A.E.Carr, M.Congreve, C.W.Murray, D.C.Rees. *Drug Discov. Today*, **10**, 987 (2005)