Instruct-ULTRA

Instruct Integrating Biology
Instruct Academic Services Ltd

Position paper on the operation of NMR services and RIs in general as a guide for Users

Executive Summary

This document reports the results of an NMR European landscape analysis carried out in the frame of the activities of the Instruct-Ultra project. The aim is to present to the broad scientific community the opportunities offered in Europe to access NMR facilities and expertise. "The biological NMR open access landscape" chapter, together with the map of "Location of European RIs offering biological NMR access to external parties" summarising NMR access opportunities at European scale, is intended to be a continuously updated guide specifically addressing non expert users, wishing to exploit the use of this technology in their research.

NMR spectroscopy in Life Sciences

Nuclear Magnetic Resonance (NMR) spectroscopy is an enabling technology capable to provide information and answers to biological problems that cannot be obtained by other means. NMR studies, both in solution and in the solid state, can inform on the structure of a macromolecule in many different environments ranging from buffered solutions to intact cells, can provide insight in dynamic processes, and allow to monitor biomolecular interactions that are key to the cellular response to environmental, developmental and growth signals (Figure 1).

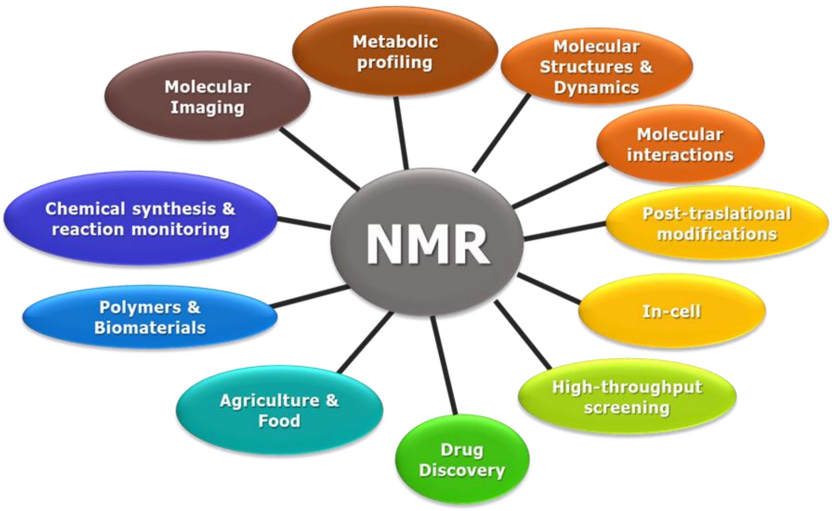

Figure 1. Areas central to biological and biomedical research where NMR has an enhancing position with respect to other methods of structural analysis

As such, NMR is central to the study of folding, unfolding and disordered states of proteins because of its capability to define the structures of proteins in solution and to characterize the

dynamic properties that are inherent to function. Understanding protein folding and the mechanisms underlying molecular assembly formation enables us to design proteins with new functions, and to characterize the kinetic pathways that control these processes *in vitro* and *in vivo*. Since folding is also necessary to generate functional nucleic acids (DNA, RNA) and in the self-assembly of complex molecular machines such as viruses, understanding the folding process is an indispensable part of modern drug discovery strategies. The interaction of small molecules (drugs) with proteins is being more and more successfully exploited by NMR, also with high-throughput approaches focusing on increasingly large ensembles of chemically related small compounds that are obtained by combinatorial chemistry.

Protein unfolding is just as crucial as their folding. Many proteins are functional without being structured and others trigger a biological response which is coupled to the folding process. Functional disorder is abundant in higher multi-cellular organisms, in particular in regulatory protein regions that orchestrate dynamic cellular functions relying on spatial and temporal malleability. NMR allows the study at atomic resolution of highly mobile macromolecules in a variety of conditions enabling the description of the different conformations that are sampled during their function.

Unfolding can also be related to misfolding events responsible of malfunctioning of living systems, and an increasing range of diseases are now associated with such problems. The generation and proliferation of protein fibrils and plaques is linked, for example, to the occurrence of amyloidosis diseases, including Parkinson’s and Alzheimer's disease. Solid-state NMR approaches permit to characterize also structural states of fibrillary proteins even in the absence of regular order.

NMR enables also to follow in real time key events and processes that are related to the function of proteins, such as the occurrence of post-translational modifications or other intermolecular interactions. Such studies can provide a better understanding of the regulatory role of particular protein modifications influencing its binding to various functional partners, a process at the heart of cellular signalling and other regulatory mechanisms.

Unique to NMR spectroscopy is the investigation of the mobility of the protein polypeptide backbone but also on their amino acid side-chains, both internally and on the surface of the protein. Not only the side chain conformation but also the dynamic interconversion between multiple side chain conformations are important structural features that are accessible to NMR.

These often play an important role in ligand binding or catalysis, in recognition processes at the protein surface, or in stabilizing native molecular structure. In solution, amino acid side chains at the surface and larger exposed loops are generally more mobile than residues in the core of the protein, and their dynamics can be characterized in detail by studies of hydrogen-deuterium exchange of amide protons and heteronuclear relaxation experiments. Other aspects which can be investigated by NMR are the ionization state of functional groups in side chains, the identification of metal ion coordinating groups, and others. NMR spectroscopy has also been shown to be most useful in the determination of the location and lifetime of hydration water in bio-macromolecules both in solution and in solid-state. Water molecules can act as integral parts in both proteins or nucleic acids, and are known to be involved in a number of important interactions: they compete for or mediate hydrogen bonds, they provide an electrostatic screening, they form an essential element of the hydrophobic effect, and they play major roles in ligand binding and molecular recognition processes.

Another key aspect which is unique to NMR is the characterization of transient interactions between or within biomolecules. In most of the functional processes partner molecules should interact transiently while the specific action is in progress. This dynamical condition can only be characterized in solution and solution NMR is in practice the only available experimental technique to provide atomic resolution information. Protein-nucleic acid interactions form also a field of research which can be addressed almost exclusively by NMR, being able to monitor complexes that involve transient/low-affinity molecular interactions. The DNA-sequence of a gene does not suffice to fully explain its function, as multiple layers of regulation at the transcriptional and post-transcriptional stage determine level of protein expression in response to environmental, developmental and growth signals. NMR is key to understand how DNA- and RNA-binding proteins interact with their cellular targets and form different higher-order macromolecular assemblies during gene expression and its regulation. There is also an increasing effort to characterize the structure and interaction of non-coding regions of mRNAs by NMR spectroscopy, since mutations that correlate with chronic disease or birth defects increasingly map to the non-coding regions of genes or to the proteins that bind them.

Finally, it is also possible to exploit the unique ability of NMR as a non-invasive *in vivo* technique to study metabolic pathways in relation to constitutive or regulated gene expression patterns. Recent advancements in in-cell NMR, i.e. the high-resolution characterization of biomolecules in living cells, has opened up the possibility to tackle structural and functional aspects in the

physiological context, maintaining the atomic resolution that is accessible by NMR. In addition, performing metabolomics by NMR is strongly surfacing, since it allows to analyse hundreds or thousands of low molecular weight metabolites in intact tissue or bio-fluids such as urine, serum, plasma, saliva and so on to report on the physiological state of an organism. Furthermore, it is establishing itself as an important technique in the field of food analysis and human nutrition in general.

The NMR user community

During the last 20 years the use of NMR in analytical chemistry and life science research has been growing steadily, with increasing impact on the European scientific community.

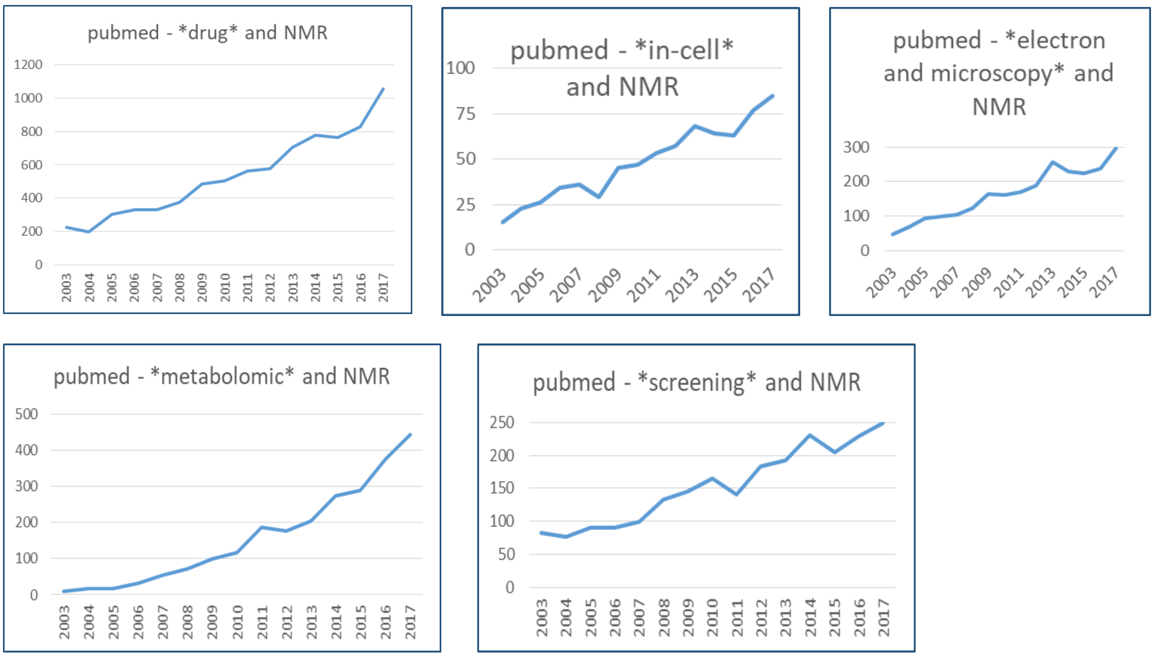

Figure 2. Numbers of scientific publications found in Pubmed (www.pubmed.com) as a function of time, when searching for 'NMR' in the context of different key words

A position paper for stakeholders (available here) produced in the frame of the EU-funded structural biology project Bio-NMR by the end of 2014 investigated this, showing steep increase in the number of peer-reviewed publications that report the use of NMR technology. This recognized trend is continued also in recent years, particularly in areas where NMR is establishing itself as a reference technique. This is for instance illustrated by the statistics in fields such as drug discovery, metabolomics and screening. It is worth mentioning that the same

tendency is found in the number of publications related to novel NMR techniques such as “in-cell NMR” as well as in the growing number of research using an integrated approach.

Not only the high-field NMR equipment is expensive, it's operation also requires dedicated housing and expert staff. To make the technique available to a broad research community, (large) NMR infrastructures have been setup that are open for access by external users. The growing interest of the wider scientific community in NMR applications is reflected in the demand for NMR access by external parties, which has been high over the years and is still growing.

At present day several initiatives are in place that allow open access to academic and industrial researchers, covering the scientific areas where NMR can be useful either as a stand-alone technology or integrated with others techniques. It is thus timely to initiate a coordinated action to optimize the use of such precious NMR resources, which can also function as a guide for the external (potential) NMR user community to discover the possibilities that best match their scientific needs. This is of foremost importance, as the external users of the equipment pinpoints the versatile demands of the interested scientific communities. In addition, any efficient and productive interaction between users and NMR Research Infrastructure (RI) scientists will strongly consolidate the use of NMR spectroscopy in the progress of science. The current European initiatives for access to NMR RIs in the field of life sciences are shortly described in the next paragraph.

The biological NMR open access landscape

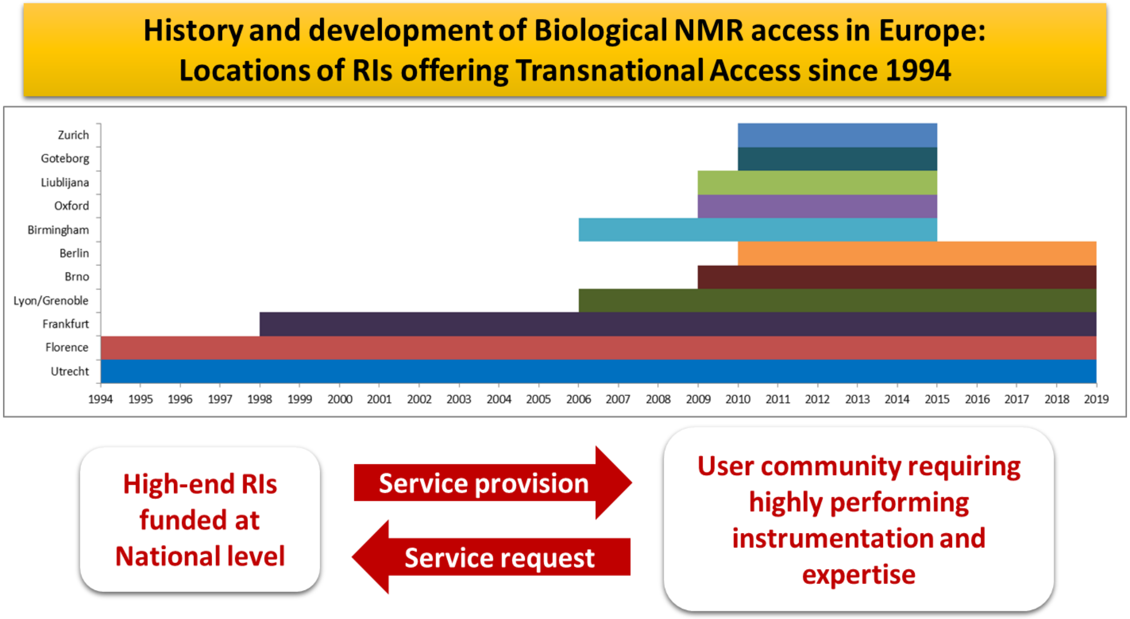

Figure 3. History of biological NMR access to external researchers in Europe

With a cumulated capital investment well above 500 Million Euro, European biological NMR Research Infrastructures (RIs) fulfil a prominent role in several distributed Research Infrastructures for Life Sciences. Since 1994, a continuously increasing grid of NMR RIs provides transnational Europe-wide open access to state-of-the-art NMR technology as a distributed network (Figure 3). The involved NMR facilities have a strong record of providing integrated access and are often considered a model with impact that is highly appreciated, even outside Europe.

Many countries in Europe have national programs of access to NMR facilities and, in many cases, these match the requirements of the users both in terms of technical support and equipment. However, often forefront research projects benefit from specific expertise and high-end instrumentation. Then, access is often sought trans-nationally and the European RIs are the reference choice. Whereas the construction of these RIs have been based on national, regional and institutional investments, the international access is organized through different international networks, repeatedly stimulated by EU-funding. To utilize the precious NMR capacity optimally, trans-national access is always based on scientific merit. This is effectuated by peer review of research proposals submitted by external users, taking into account primarily

the scientific quality of the proposed research and the feasibility and purpose of the experiments. Users may be scientists with or without previous experience in NMR, as the RIs may provide, beside technical assistance for the acquisition of the data, highly qualified scientific expertise for analysis and support.

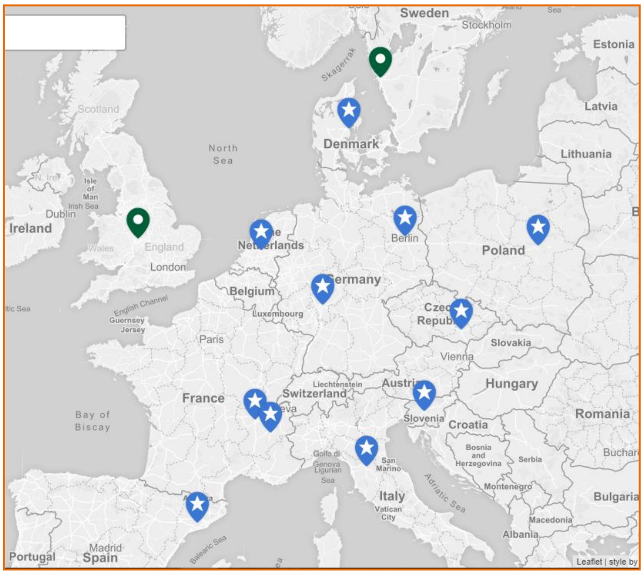

Figure 4. Location of European RIs offering biological NMR access to external parties

The map beside (link) shows the Research Infrastructures that offer access and/or expertise at European level in the field of biological NMR spectroscopy and beyond. All these RIs couple high-end NMR instrumentation to a long-standing experience of access provision and the capacity of supporting also non-expert users.

The indicated NMR facilities provide NMR access through one or more of the access routes available at European level described here below. Users can choose the route that fits best with their necessities. Users that are not sure of which may be their best access route are invited to contact NMR scientists at these NMR RIs.

INSTRUCT-ERIC, landmark of the ESFRI (European Strategy Forum on Research Infrastructures) Roadmap, is a distributed European Research Infrastructure facilitating user access to cutting-edge technology in structural biology, scientific expertise and offering specialist training. INSTRUCT-ERIC promotes research projects that demonstrate innovative approaches within INtegrative STRUCTural biology and mediates access to high-end, specialist instrumentation that is

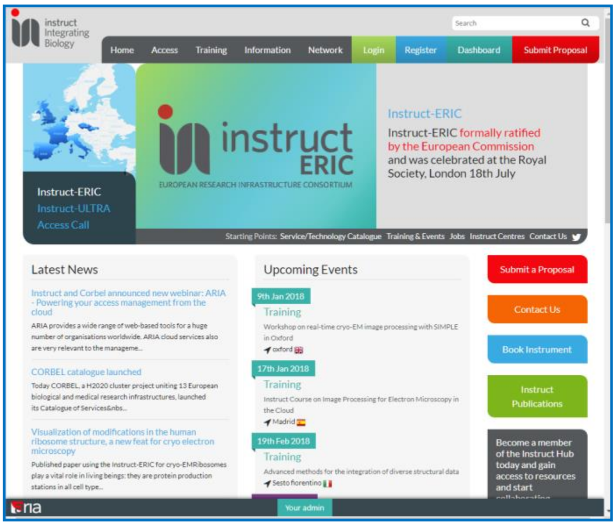

Figure 5. The Instruct-ERIC website

available at different INSTRUCT-ERIC Centres. Embracing the entire area of structural biology, the catalogue of available techniques is very wide and can be consulted at the INSTRUCT-ERIC homepage (www.structuralbiology.eu, Figure 5).

INSTRUCT-ERIC Centres currently include four NMR platforms: CERM/CIRMMP (Firenze, IT), UU (Utrecht, NL), RALF-NMR (Grenoble, FR) and CEITEC (Instruct-CZ, Brno, CZ) of which the first two specialize also in solid-state NMR. Researchers from INSTRUCT-ERIC member states (www.structuralbiology.eu/content/countries-instruct) are eligible for financial support covering (part of the) access costs and travel expenses to the facility.

User applications for access can be submitted at any time, although also periodically special calls for access are announced with a defined deadline. Even though the long-term aim of INSTRUCT-ERIC is to encourage the integrative use of technology and methodologies, applications for stand-alone technology such as NMR are also possible.

iNEXT (Infrastructure for NMR, EM and X-rays for Translational research) is a EU-funded H2020-INFRAIA project offering transnational access to study the structure and function of biological macromolecules and their assemblies, with the specific aim of translating fundamental research into bio-scientific applications. iNEXT (www.inext-eu.org) allows open access for researchers from academia and industry alike to services in X-ray crystallography, Small Angle X-ray Scattering (SAXS), NMR, Electron Microscopy (EM), light imaging and other biophysics techniques for instance to study macromolecular interactions. Through iNEXT, high-end instrumentation both for solution and solid state NMR is available at BMRZ (Frankfurt am Main, DE), CERM/CIRMMP (Florence, IT), FMP-NMR (Berlin, DE), RALF-NMR (Grenoble/Lyon, FR) and UU (Utrecht, NL). BMRZ, CERM/CIRMMP and UU also accept so-called “enhanced support” projects, which can include training in the use of the instruments

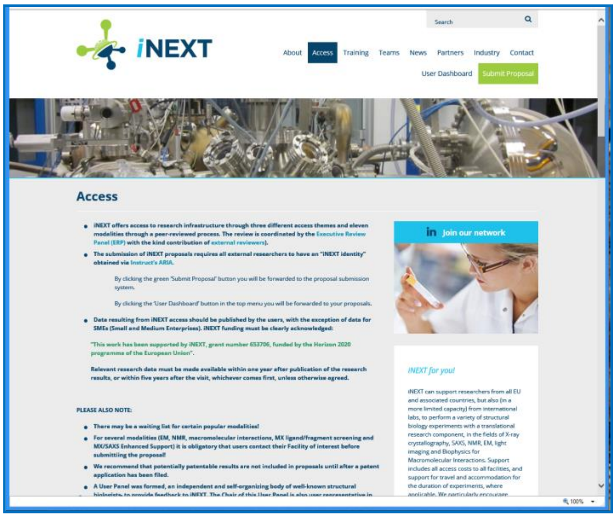

Figure 6. The iNEXT website

and in the analysis of the data, and/or preliminary measurements to evaluate whether samples are suitable for future NMR approaches.

iNEXT support covers the costs for access at all the RIs involved, in many occasions also travel and living expenses for researchers not only from all EU and associated countries, but also (up to a maximum) from non-EU laboratories.

EuroBioNMR European Economic Interest Grouping (EEIG) is a coordinated organization of NMR facilities that has recently been established to facilitate and reinforce interactions between NMR RIs and research groups wishing to introduce NMR approaches in their research. EuroBioNMR (www.eurobionmr.eu) offers collaborative access to the broad life science community including chemical, pharmaceutical, food industries and agricultural sciences in addition to structural biology services. At present EuroBioNMR includes CIRMMP (Florence, IT),

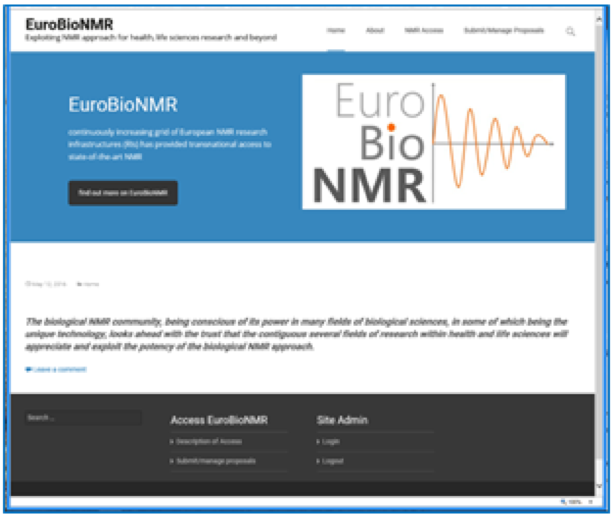

Figure 7. The EuroBioNMR EEIG website

BMRZ (Frankfurt, DE), the Bijvoet Center (Utrecht, NL), the Slovenian NMR Centre (Ljubljana, SI) the Laboratory of New Methods of NMR Spectroscopy (Warsaw, PL), the Danish Center for Ultrahigh-Field NMR Spectroscopy (Aarhus, DK) and the University of Barcelona (Barcelona, ES). These seven NMR centres cover a wide range of applications which are going beyond structural biology, with expertise in most research fields within the health and life sciences, not excluding possible applications in material science and other fields.

CORBEL (<u>CO</u>ordinated <u>R</u>esearch Infrastructures <u>B</u>uilding <u>E</u>nduring <u>L</u>ife-science Services) is an EU-funded H2020 project that provides pilot access to resources and services that enable cutting-edge biomedical research at the heart of the participating ESFRI RIs within the health and food domain. Currently there are 13 biological and medical RIs in Europe, joining forces in CORBEL (www.corbel-project.eu) towards offering shared services for life-science. All involved RIs offer services that allow researchers from academia and industry to utilize their facilities, technologies and expertise. Services can simply be consultation by experts, but also access to data and biological samples, the use of data analysis tools, the access to instrumental facilities plus support from technicians, and much more.

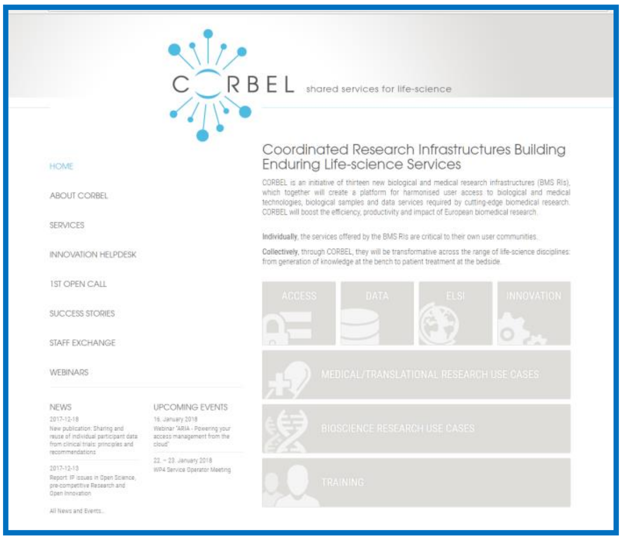

Figure 8. The corbel website

EXCEMET (<u>EX</u>pert <u>CE</u>nter in <u>MET</u>abolomics) combines experimental/analytical and informatics/computational expertise. The EXCEMET goal is to strengthen interactions between the metabolomics research community and biobanks, as well as to help the biobanks to develop increasingly higher standards for their sample quality. EXCEMET (www.excemet.org) complements existing activities of the biobanks, extending access opportunities from high quality samples to advanced biomolecular analyses, and the dissemination of curated knowledge from such analyses in open-access, long-term maintained databases.

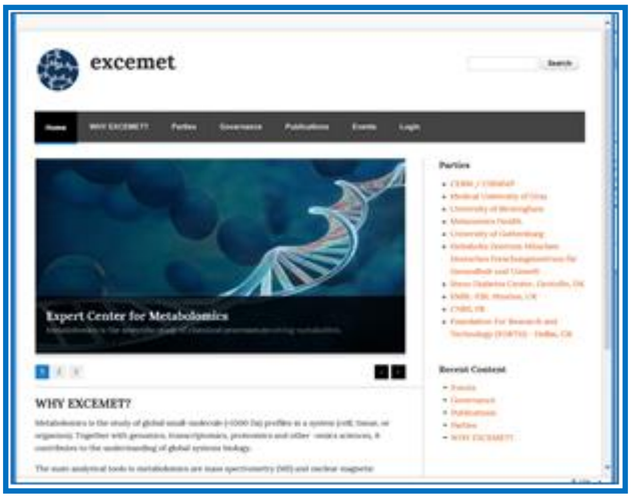

Figure 9. The Excemet website

Presently, the main analytical tools used in metabolomics are mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. The NMR partners of EXCEMET (CERM/CIRMMP, University of Birmingham, University of Gothenburg and CNRS-FR) have the analytical competences as well as the IT and computational expertise for data analysis, biomedical data interpretation, data visualization and ontology definition. Even though EXCEMET partners do not offer access to their instrumentation, they can be considered a reference for the provision of expertise for tailored applications and for quality assurance in metabolomics studies.

West-Life (a Virtual Research Environment for Structural Biology) is an EU-funded H2020-EINFRA project that provides services for computation and data management to researchers in structural biology, integrating multiple approaches and experimental techniques, including NMR. West-Life builds on European e-Infrastructure solutions developed by EGI and EUDAT and connects web services with repositories for structural biology. Services, support and tutorials are available at https://about.west-life.eu.

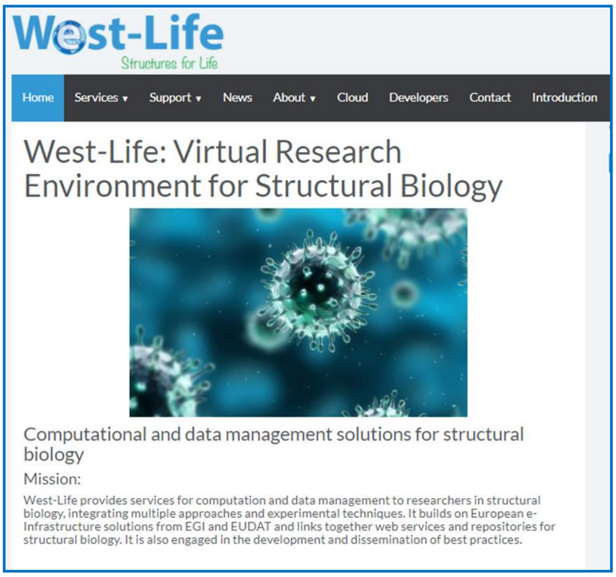

Figure 10. The West-Life website

Contact us for further information!

admin@structuralbiology.eu

info@inext-eu.org

info@eurobionmr.eu