

e-NMR gLite grid enabled infrastructure

Nuno Loureiro-Ferreira¹, Tsjerk A. Wassenaar¹, Sjoerd J. de Vries¹, Marc van Dijk¹, Gijs van der Schot¹, Johan van der Zwan¹, Rolf Boelens¹, Andrea Giachetti², Dario Carotenuto², Antonio Rosato², Ivano Bertini², Torsten Herrmann³, Anurag Bagaria⁴, Victor Zharavin⁴, Hendrik R.A. Jonker⁴, Peter Güntert⁴, Harald Schwalbe⁴, Wim F. Vranken⁵, Stefano Dal Pra⁶, Mirco Mazzucato⁶, Eric Frizziero⁶, Sergio Traldi⁶, Marco Verlato⁶, Alexandre M.J.J. Bonvin¹

¹ Bijvoet Center for Biomolecular Research, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH, Utrecht, The Netherlands.

{n.l.ferreira,a.m.j.j.bonvin}@uu.nl

² Magnetic Resonance Center, University of Florence, 50019 Sesto Fiorentino, Italy.

³ University of Lyon, Centre de RMN à Très Hauts Champs, 69100 Villeurbanne, France.

⁴ Institute of Organic Chemistry and Chemical Biology and Biomolecular Magnetic Resonance Center, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany.

⁵ European Bioinformatics Institute, Hinxton, Cambridge, CB10 1SD, United Kingdom.

⁶ Istituto Nazionale di Fisica Nucleare, Sez. di Padova, 35131 Padova, Italy.

Abstract. The e-NMR project is an European e-infrastructure that aims at providing the bio-NMR community with a software platform integrating and streamlining computational approaches necessary for NMR data analysis. The infrastructure is grid enabled with fifteen gLite based partners sharing computational resources. A main focus of the consortium is to provide protocolled services through easy-to-use web interfaces, while retaining sufficient flexibility to handle specific requests by expert users. Various programs relevant for structural biology scientists are grid ported and already available through the e-NMR web portal, including HADDOCK, XPLOR-NIH, CYANA and CS-ROSETTA among others. A general overview of the project current status toward EGEE/EGI integration, as well as brief guidelines on how to become an e-NMR site/user will be considered. With more than 170 registered users, enmr.eu is currently the second largest virtual organization in the life sciences. The state of the project can be found on the web page <http://www.enmr.eu>.

1 Introduction

Nuclear Magnetic Resonance (NMR) spectroscopy has established itself as an invaluable tool to investigate the structure, dynamics and molecular interactions of biological macromolecules, such as proteins [2]. This knowledge is of paramount importance for studying the functions and mechanisms of action of macromolecules *in vivo*. Design strategies for rational drug design is strongly coupled to the need of having three-dimensional molecular models describing the intervening partners.

NMR plays an important role in life sciences at both European and international levels. Two Integrated Infrastructure Initiatives (I3) are operative in the field, providing users access to state-of-the-art NMR instrumentation as well as technical expertise (EAST-NMR and EU-NMR, [35,38]). Additionally, the European Commission (EC) has funded a coordination action aimed at the establishment of common experimental approaches and at the spreading of best experimental practices across Europe (NMR-Life, [44]). Altogether, these initiatives provide a reference point for the large majority of European scientists with an interest in bio-NMR. At the same time, European developments in network research infrastructures in the past years resulted in a leading edge high-speed network covering all Europe. A production grid infrastructure was set up by projects such as EGEE [37], providing a 24/7 based platform for global collaborative research – e-Science.

The standard data processing in NMR structure determination of proteins consists of a sequence of successive steps [23]: i) Fourier transformation of the time-domain data to obtain the frequency-domain spectra, ii) assigning the NMR signals to individual atoms in the polypeptide chain, iii) collecting structural restraints, iv) using structural restraints to determine atomic coordinates, and finally v) validating the obtained structure models. This rather lengthy process is computational demanding, usually requiring powerful in-house computers and data storage, and a set of programs not always easy to handle. These factors have thrown up a barrier for many research groups in life sciences to employ the full power of NMR.

Automating as many as possible of these steps would transform NMR structure determination of macromolecules, into a highly reproducible activity, decoupling the subjective decisions and expertise of an individual to well established protocols [7]. The e-NMR project started as an European grid initiative to facilitate data processing, both to allow groups lacking the resources to add NMR to their toolbox, as well as to allow dedicated NMR groups to shift their attention from standard tasks toward cutting-edge research.

2 VO identity card

The e-NMR project is EC funded, aiming to optimize and extend the use of the NMR research infrastructures of EU-NMR, through the implementation of an e-infrastructure. The bio-NMR user community would thus have a platform able to integrate and streamline computational approaches required for bio-NMR data analysis. The e-NMR infrastructure is based on the European grid infrastructure. A global scope virtual organization (VO) named `enmr.eu` was created under the life sciences discipline, and its VO identity card is registered at the EGEE Grid Operations Centre (GOC) [49]. All grid sites supporting the `enmr.eu` VO are being coordinated by the respective Regional Operations Centre within each regional federation, and/or the corresponding National Grid Initiatives (NGI). Currently, the e-NMR project involves an operational grid, on which various NMR related software packages have been and are being grid ported. A specific aspect of the

project is the implementation of easy to use web interfaces while retaining sufficient flexibility for expert users.

3 Grid deployment status

The main partners of the e-NMR project represent the three core members of the existing European NMR I3 EU-NMR, namely the Biomolekulares Magnetresonanz Zentrum der Universität Frankfurt (BMRZ), the Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine Paramagnetiche (CIRMMP), and the Bijvoet Center for Biomolecular Research (BCBR). Together, they provide a large amount of expertise regarding consolidation and wide spread of NMR good practices, both in practical and computational aspects. e-Science expertise is provided by Istituto Nazionale di Fisica Nuclear (INFN), supporting the deployment of grid middleware and providing the link to the EGEE project.

The e-NMR infrastructure is based on gLite, a lightweight middleware for grid computation [26]. gLite provides a framework for building grid applications tapping into the power of distributed computing and storage resources across the Internet. The main project partners deployed the grid profiles using a gLite version customized and distributed by INFN [30]. The INFN Release is 100% compatible with gLite 3.x, with some additional advanced accounting and monitoring features of interest to support both regional and international VO's in Italy. It also allows to automatically enable the enmr.eu VO at a grid site. See section 6 for further details. Currently, 15 grid sites are supporting this infrastructure, either running gLite INFN release or the "pure" EGEE gLite version. An overview of all sites currently supporting the project can be assessed from GStat2.0 web portal [29].

The e-NMR grid comprises three resource centers (BCBR, BMRZ, CIRMMP) fully dedicated to the enmr.eu VO, with high priority queues specific for VO operational testings. Since the beginning of the project, e-NMR also includes a set of resource centers part of the EGEE production grid which agreed to support the project, namely sites managed by the Italian Grid Infrastructure [42] and the Dutch BiG Grid [34]. More recently, the South African National Grid [45] and the UK National Grid Service [43] decided to support enmr.eu on a best effort basis, sharing their resources for the benefit of their national bio-NMR community.

3.1 Grid services

Currently, central core grid services span all the areas of the gLite architecture, namely job management, information system, security, data management, and accounting. Job management services include the Workload Management System (WMS, [53]), responsible for distributing and managing of tasks across grid resources, and the Logging and Bookkeeping [25], responsible for tracking WMS jobs as they are processed by individual grid components. Information system services include several top level Berkeley Database Information Indexes (BDII, [33]), providing detailed information about grid services which is needed for various different tasks. Security services deployed at e-NMR infrastructure include Virtual Organization Membership Services (VOMS, [39]) and several Hydra servers [55]. The

	CPU ^s *	KSI2K [†]	Storage [‡]
Total dedication	252	661	3.34
Total shared	9449	16265	147.43
Sum	9701	16926	150.77

* Number of cores; [†] KSI2K unit is used to benchmark CPU performance for integer operations [46]; [‡] In TB units.

Table 1. e-NMR computing and storage resources.

VOMS serves as a central repository for user authorization information, whereas Hydra is a file encryption storage solution. A centralized LCG File Catalog [27] is being used as a data catalog containing logical to physical file mappings. Several VO monitoring services, namely Gstat, SAM, and WMSMonitor, are being used to assess and monitor the health status of all sites supporting the VO. And finally, a Distributed Grid Accounting System (DGAS, [47]) accumulates information about the usage of grid resources by the users and applications. Some of these services have a broad scope inside EGEE, others are self-contained for the e-NMR project.

Along with the core services briefly enumerated above, computing and storage resources are part of the e-NMR infrastructure, either fully dedicated to enmr.eu or shared among other regional and global VO's. Table 1 discriminates those resources. Notably, 150 TB of disk space and almost 10k CPU cores are available to the bio-NMR community.

3.2 Integration with EGEE/EGI

One of the goals of e-NMR is to be fully integrated into the EGEE production grid, as well as preparing the future evolution of EGEE toward the European Grid Infrastructure (EGI) model. EGI will be the coordinator of the European grid, and is based on the federation of individual NGI's, to support a multi-disciplinary user community. Most of the deployment and operational tools designed for the EGEE and the Italian IGI production grid infrastructures have been successfully adopted by the e-NMR grid, since the beginning of the project. Furthermore, the partnerships established with EGEE and especially with the NGI's, put e-NMR in the position of managing smoothly the transition to EGI. In fact, most of the sites supporting us are already certified, while a minority are in the process of being certified or added to the GOC database [54]. Operational support complementary to EGEE official procedures is anyway provided by e-NMR teams for helping new sites willing to join e-NMR grid, even if not yet affiliated with EGEE.

4 The e-NMR web portals

A grid without users is a nonsense effort. e-NMR is proud to dedicate most of its human resources building well established practice protocols online for processing NMR data, including most of the steps for structure determination described in

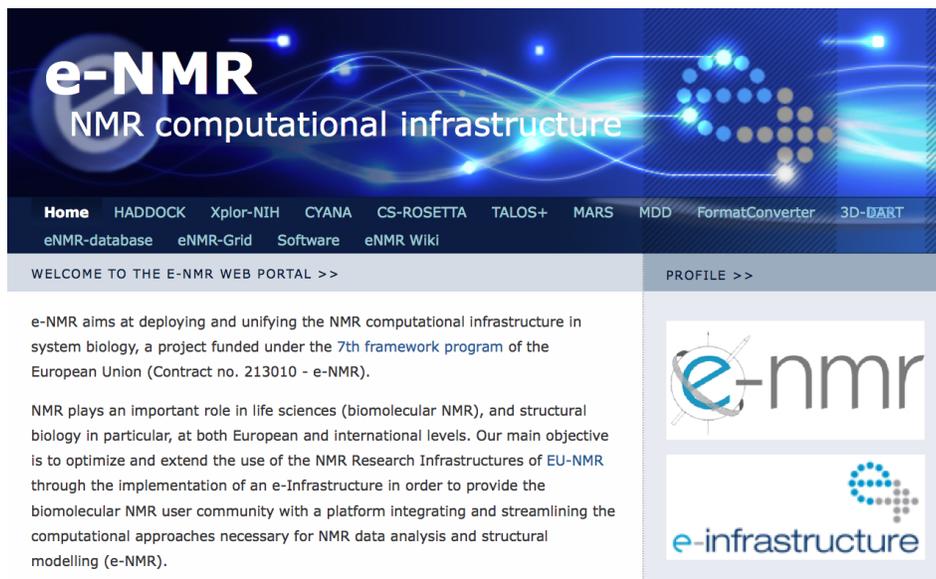


Fig. 1. The e-NMR web portal.

the introduction. A full platform for NMR based structural biology can thus be accessed via the web, taking out from the hands of the user some cumbersome aspects such as installation of software. At the same time we allow data intensive protocols to be run on the grid, which would be otherwise impossible on commodity computers. To achieve this goal, the enmr.eu developers team started by implementing individual programs as web based servers. The programs were prioritized on the basis of their popularity among NMR spectroscopists, and CPU demands, as well as usefulness for the project itself. Currently, nine portals are operational and can be accessed at <http://www.enmr.eu/webportal/> (Figure 1).

The main programs and their use within the e-NMR portal were already described [19]. Briefly, MDDNMR for non-uniformly sampled n-D spectrum processing using multi-dimensional decomposition [12], automatic backbone assignment of $^{13}\text{C}/^{15}\text{N}$ labeled proteins can be accomplished with MARS web portal [24], and TALOS+ predicts protein backbone torsion angles from NMR chemical shifts [15]. Moreover, four portals allow users to calculate the structure of proteins and/or protein complexes from NMR data, namely CS-ROSETTA [16,17], CYANA [9,11], HADDOCK [5,21], and XPLOR-NIH [14]. The latest added web portal, 3D-DART [4,13], generates nucleic acid structure models, which can later be used as starting templates during NMR structure calculations involving such biomolecules. In addition, the FormatConverter developed within the CcpNmr initiative is available for NMR related data format conversion [20]. It provides interoperability of the programs involved in NMR data processing.

The various programs ported to the e-NMR grid have different complexities during their setup, namely the number of user defined parameters describing the

algorithm. Nonetheless, the portals offer almost full control on the parameters. Full functionality of these programs through the web portal requires putting forth a complex form, contrasting with the objective of having a simple interface. To avoid compromises regarding user friendliness and functionality, two innovations were introduced in the design of the portal. The portal is divided in different interfaces, corresponding to different levels of control and user experience. At the same time, to facilitate users task and keep the forms manageable, foldable menus were introduced that group related parameters under a single header.

The use of the web portals is very straightforward and intuitive involving two steps: authentication to the portal, and program parameters setup. Albeit portals were designed and implemented following unified guidelines, some key elements do differentiate them, mainly regarding user interaction. All portals describe their usage prerequisites, along with data test cases available for the general community. Nonetheless, registration and/or access for most of the web portals requires from the user an authentication with a grid enabled e-NMR certificate. Further details, to the end user, on how to register to the grid and to the enmr.eu VO will be given in section 6.2. Once authenticated to the system, the user can submit jobs to the grid by filling in the form where parameter setup and data upload fields are present. He is notified on the status of his/her job either by e-mail or browsing the job link generated upon submission.

Successfully running web portals requires a proper machinery to handle the requests. Three layers of operation were defined in the design strategy. A server level involves handling of service requests, either by direct human interaction or through requests from another machine. This stage includes input type checking. The next level involves preparation, trafficking and monitoring of jobs between the servers and the grid. The third layer involves the process(es) to be run on the worker nodes. This component based development approach, has the advantage that such building blocks can be easily maintained, adapted or reused. A description of the implementation details behind the web portals was described in two recent publications [19,22].

5 VO resource usage

In April 2010, the VO reached the mark of 173 members, making us the second largest VO in the life sciences grid discipline. An overview of current worldwide distribution is presented in figure 2, mainly concentrated in European countries. The number of users has been increasing steeply as a result of several e-NMR workshops held in the last year. The availability of applications accessible via web portal also increased at the same time. Currently, eighteen applications were successfully ported to the Grid, half of which can be accessed through a web portal (summarily referred in sections 4 and 6.2).

The DGAS [50] middleware profile has been used together with the EGEE Accounting Portal [48], to account the usage of the grid by enmr.eu VO members. Since the official opening of the e-NMR grid (1 of November 2008), about

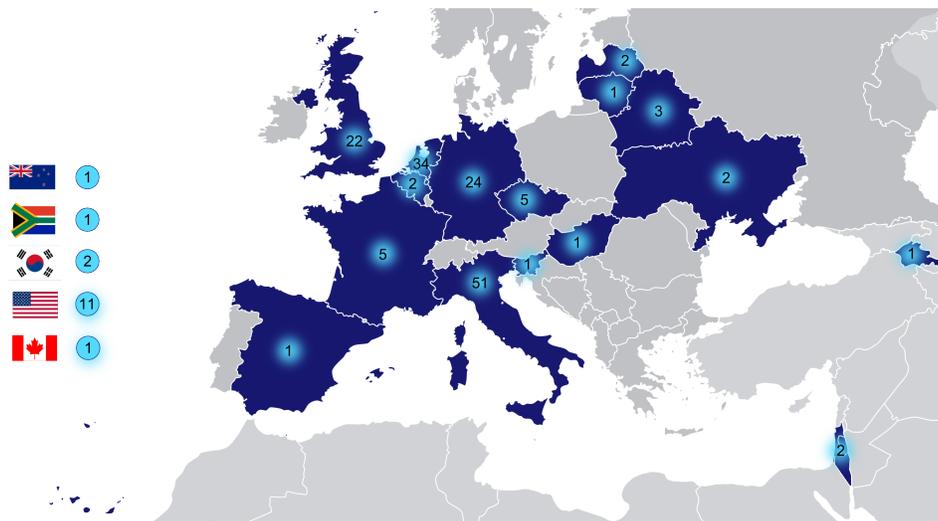


Fig. 2. Geographical view of e-NMR users distribution.

307 KSI2K.years of normalized CPU time* was consumed, with a grand total of 474K jobs submitted, representing an average of 28K jobs per month. In January 2010 a *per* application accounting mechanism was put in production, leveraging the granularity made available by the VOMS service. The most active application-s/portals for the time period since January are CS-ROSETTA and HADDOCK with 4.6K and 2.4K jobs, respectively. The number of registered users using the available web-portals varies from application to application, with for example 47 users using actively CS-ROSETTA and 89 users using XPLORE-NIH.

6 How to become enmr.eu aware

6.1 Site level

Configuring a gLite based site to support enmr.eu VO is a relatively easy task. Briefly, it involves the following four steps well known to site grid administrators:

1. Preparation of configuration files
2. Configure the VO software area
3. Verification of the installation
4. Configure the middleware services

The first step requires the preparation of three configuration files, namely the main configuration file *your-site-info.def*, the pool account user mapping *your-users.conf*, and the VOMS group mapping *your-groups.conf*. Some VO related

*The normalized CPU time is the clock time consumed on the computing nodes in KSI2K.years units.

variables should be edited in the main configuration file, for completeness, the list of VO's supported (add enmr.eu to *VOS*) and the queues made available to enmr.eu (enable at least one *queue-name_GROUP_ENABLE* queue to the VO). Other specific VO variables can be added either directly to *your-site-info.def* file or to *vo.d/enmr.eu* file, inside your site configuration directory (this is the default way in enmr.eu, see listing 1.1). Spaces and breaking lines are outlined with a symbol `_` for readability.

```

1 SW_DIR=$VO_SW_DIR/enmr
2 DEFAULT_SE=$SE_HOST
3 STORAGE_DIR=$CLASSIC_STORAGE_DIR/enmr
4 VOMS_SERVERS=" 'vomss://voms2.cnaf.infn.it:8443/voms/enmr.eu?/enmr.eu' \_
   'vomss://voms-02.pd.infn.it:8443/voms/enmr.eu?/enmr.eu' "
5 VOMSES=" 'enmr.eu\_voms2.cnaf.infn.it\_15014\_
   /C=IT/O=INFN/OU=Host/L=CNAF/CN=voms2.cnaf.infn.it\_enmr.eu' \_ 'enmr.eu\_
   voms-02.pd.infn.it\_15014\_
   /C=IT/O=INFN/OU=Host/L=Padova/CN=voms-02.pd.infn.it\_enmr.eu' "
6 VOMS_CA_DN=" '/C=IT/O=INFN/CN=INFN\_CA' \_ '/C=IT/O=INFN/CN=INFN\_CA' "

```

Listing 1.1. Enabling the enmr.eu VO in your site gLite services.

Regarding pool account user mapping and VOMS group mapping, add to *your-users.conf* and *your-groups.conf* files respectively, users and groups for the enmr.eu VO which fit your site's policy in users management. A simple example can be found in listing 1.2 where one single user is created, either to be mapped to a regular user, to a general VO software manager account, or to local VO software manager for the main three NMR sites deploying software (BCBR, BMRZ, CIRMMP).

```

1 46001:enmr001:46000:enmr:enmr.eu::
2 ...
3 46901:sgmenmr001:46090,46000:sgmenmr,enmr:enmr.eu:sgm:
4 ...
5 46921:sgmenmrbcb001:46091,46000:sgmenmrbcb,enmr:enmr.eu:sgmbcb:
6 ...
7 46941:sgmenmrbrz001:46092,46000:sgmenmrbrz,enmr:enmr.eu:sgmbrz:
8 ...
9 46961:sgmenmrcirmmp001:46093,46000:sgmenmrcirmmp,enmr:enmr.eu:sgmrcirmmp:
10 ...

```

Listing 1.2. enmr.eu pool account user mapping.

A suitable group configuration file for the VO should contain the following entries in listing 1.3.

```

1 "/enmr.eu/ROLE=SoftwareManager" :::sgm:
2 "/enmr.eu/bcbr/ROLE=SoftwareManager" :::sgmbcb:
3 "/enmr.eu/bmrz/ROLE=SoftwareManager" :::sgmbrz:

```

```
4 "/enmr.eu/cirmmp/ROLE=SoftwareManager" :::sgmcirmmp:
5 "/enmr.eu/*" ::::
6 "/enmr.eu" ::::
```

Listing 1.3. enmr.eu VOMS group mapping.

The second stage of enabling the VO deals with the software area exported to the computing element(s) and attached worker nodes. A specific structure is needed for software installation based on sgm pool accounts, as can be deduced from the user and group mapping definitions above. To fit these requirements, some additional steps have to be performed in the software area defined by system variable *VO_ENMR_EU_SW_DIR*. Assuming that this software directory is already created with *sgmenmr001.sgmenmr* ownership, a script similar to the one listed in 1.4 has to be issued in order to create the VO software pools with correct permissions. This area can be read by any member of the VO, but can only be written by the software manager of the VO. This method avoids the overheads with the repeated download and installation of software packages, often encountered in the other software deployment methods (sandbox and parachute). Currently, a minimum of 10 GB of storage space is required by enmr.eu.

```
1 #!/bin/bash
2 for SITE in BCBR BMRZ CIRMMP ; do
3   site=`echo $SITE | tr "A-z" "a-z"`
4   mkdir ${VO_ENMR_EU_SW_DIR}/${SITE}
5   chown sgmenmr${site}001.sgmenmr${site} ${VO_ENMR_EU_SW_DIR}/${SITE}
6   chmod g+w ${VO_ENMR_EU_SW_DIR}/${SITE}
7 done
```

Listing 1.4. Configure software area

If your site is purely LCG/gLite based, you will need to install the enmr.eu VOMS server certificate, which is not included in the gLite official RPM's. You can download the latest released certificate from INFN production repositories [28].

The final step is a configuration of all required nodes to be shared with the enmr.eu community. After a successful configuration, the VOMS mapping file */etc/grid-security/grid-mapfile* should contain the lines in listing 1.5.

```
1 ...
2 "/enmr.eu/Role=SoftwareManager/Capability=NULL"┆.sgmenmr
3 "/enmr.eu/Role=SoftwareManager"┆.sgmenmr
4 "/enmr.eu/bcbr/Role=SoftwareManager/Capability=NULL"┆.sgmenmrbcbr
5 "/enmr.eu/bcbr/Role=SoftwareManager"┆.sgmenmrbcbr
6 "/enmr.eu/bmrz/Role=SoftwareManager/Capability=NULL"┆.sgmenmrbmrz
7 "/enmr.eu/bmrz/Role=SoftwareManager"┆.sgmenmrbmrz
8 "/enmr.eu/cirmmp/Role=SoftwareManager/Capability=NULL"┆.sgmenmrcirmmp
9 "/enmr.eu/cirmmp/Role=SoftwareManager"┆.sgmenmrcirmmp
10 "/enmr.eu/*/Role=NULL/Capability=NULL"┆.enmr
11 "/enmr.eu/*"┆.enmr
```

```
12 "/enmr.eu/Role=NULL/Capability=NULL"_.enmr
13 "/enmr.eu"_.enmr
14 ...
```

Listing 1.5. VOMS mapping

After configuration of the relevant gLite profiles and a shared area for installation of application-specific software is in place, the enmr.eu VO software administrator will assume the responsibility to fully manage the VO specific software. Once this stage is completed, the LDAP address of the site-level BDII of the new site will be added to all enmr.eu top-level BDII's, making enmr.eu WMS's aware of the new resources.

If you are using the INFNGRID release, the task of enabling enmr.eu is simplified since example configuration files are in place, and the VOMS certificate is already installed. For further details about the steps considered above please refer to the following links [31,51].

6.2 User layer

Registering with e-NMR is less straightforward than for many other web-based services of use in the life sciences. The steps to be taken are outlined below. Whereas the first two steps are mandatory, the last step is only needed by expert users who want to make use of the gLite User Interface (UI) service.

1. Own a personal certificate issued by a IGTF recognized CA
2. Register with the enmr.eu VOMS
3. Install and configure a PnP gLite UI

The procedure for obtaining a certificate can vary from country to country. Briefly, it involves identifying oneself to a Registration Authority (RA), often located within universities or research institutions. The RA will transmit the request to the national Certification Authority (CA), which will then process it and release the certificate. The user will be contacted by the CA, with instructions on how to download and install the certificate to his own machine. Detailed instructions and web links to various European CA's are available at <http://www.enmr.eu/eNMR-registration>. Users from the United States can request their certificate directly to the Structural Biology Grid [32]. Other international users should visit the website of the International Grid Trust Federation [41].

Once in the possession of a grid certificate, and before using the computational resources available to a specific VO, users must register themselves to a VOMS supporting the VO. The e-NMR membership can be requested by filling in an online form [57]. Within a few hours the e-NMR VO administrator will approve your request.

Though the e-NMR project is web-portal centric, expert users can also make usage of the VO resources through a Plug & Play User Interface (PnP UI) pre-configured for enmr.eu. This involves the installation of software in an enabled computer compliant with gLite requisites, along with the installation of the user certificate. For further details, a user guide for end-users was made available [52]. Notice

that grid sites usually also have an UI installed, and its usage is granted to users on a per case basis after approval by the site grid administrator. A number of use-cases for use with the UI are available at <http://www.enmr.eu/eNMR-tutorials>. They provide simple examples on how to run structure calculations on the e-NMR grid with CNS [3], Xplor-NIH and CYANA, NMR data processing with PROSA [8], automated assignment with MAPPER [10], MARS and GARANT [1], scalar coupling constant determination with INFIT [18], processing of non-uniformly sampled data with MDDNMR, and molecular dynamics simulations with GROMACS package suite [40].

7 Conclusion

Driven by increasingly complex problems and propelled by new technologies, today's science is as much based on computation, data analysis, and collaboration between *in vivo* and *in silico* scientists [6]. The grid concept motivated by distributed computing infrastructures for advanced science is emerging. Coordinated resource sharing and problem solving is becoming truly a multi-institutional task.

The e-NMR consortium is making an endeavor to embrace e-science opportunities. We managed to set up an operational grid, ported eighteen NMR applications and bring up nine web portals, with several others underway. At the time of writing, e-NMR has already grown to be the second largest virtual organization within the life sciences. This successful start has been underlined by the award for the best demonstration of an application in the last EGEE 2009 user forum. The project is evolving into a factor of importance within structural biology, and life sciences in general. The final stage of the project will combine all these portal modules, providing a comprehensive set of tools for integrated analysis of NMR data, and more.

Up-to-date information regarding the state of the project, the available services, and how to join this continuing effort, can be found on the project web page at <http://www.enmr.eu>.

Acknowledgements

The e-NMR project is funded by the European Commission under an FP7 e-infrastructure grant, contract no. 213010. The national grid initiatives from Italy, the United Kingdom, Germany, and the Dutch BiG Grid project with financial support from the Netherlands Organization for Scientific Research (NWO) are acknowledged for the use of the computing and storage facilities.

References

1. C. Bartels, P. Güntert, M. Billeter, and K. Wüthrich. "GARANT - a General Algorithm for Resonance Assignment of Multidimensional Nuclear Magnetic Resonance Spectra.", *J. Comp. Chem.*, **18**, pp. 139–149, (1997).

2. M. Billeter, G. Wagner, and K. Wüthrich. "Solution NMR structure determination of proteins revisited.", *J. Biomol. NMR*, **42**, pp. 155–158, (2008).
3. A.T. Brünger. "Version 1.2 of the Crystallography and NMR system.", *Nature Protocols*, **2**, pp. 2728–2733, (2007).
4. M. van Dijk, A.D. van Dijk, V. Hsu, R. Boelens, and A.M.J.J. Bonvin. "Information-driven protein-DNA docking using HADDOCK: it is a matter of flexibility", *Nucleic Acids res.*, **34**, pp. 3317–3325, (2006).
5. C. Dominguez, R. Boelens, and A.M.J.J. Bonvin. "HADDOCK: A protein-protein docking approach based on biochemical or biophysical information", *J. Am. Chem. Soc.*, **125**, pp. 1731–1737, (2003).
6. I. Foster. "The Grid: a new infrastructure for the 21st century science", *Wiley Series in Communications Networking & Distributed Systems*, John Wiley & Sons Ltd, chapter **2**, pp. 51–63, (2004).
7. P. Güntert. "Automated structure determination from NMR spectra.", *Eur. Biophys. J.*, **38**, pp. 129–143, (2009).
8. P. Güntert, V. Dötsch, G. Wider, and K. Wüthrich. "Processing of multi-dimensional NMR data with the new software PROSA", *J. Biomol. NMR*, **2**, pp. 619–629, (1992).
9. P. Güntert, C. Mumenthaler, and K. Wüthrich. "Torsion angle dynamics for NMR structure calculation with the new program DYANA", *J. Mol. Biol.*, **273**, pp. 283–298, (1997).
10. P. Güntert, M. Salzmann, D. Braun, and K. Wüthrich. "Sequence-specific NMR assignment of proteins by global fragment mapping with the program MAPPER", *J. Biomol. NMR*, **18**, pp. 129–137, (2000).
11. T. Herrmann, P. Güntert, and K. Wüthrich. "Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA", *J. Mol. Biol.*, **319**, pp. 209–227, (2002).
12. V.A. Jaravine, A.V. Zhuravleva, P. Permi, I. Ibraghimov, and V.Y. Orekhov. "Hyperdimensional NMR Spectroscopy with Nonlinear Sampling", *J. Am. Chem. Soc.*, **130**, pp. 3927–3936, (2008).
13. X.J. Lu, and W.K. Olson. "3DNA: a software package for the analysis rebuilding and visualization of three-dimensional nucleic acid structures", *Nucleic Acids res.*, **31**, pp. 5108–5121, (2003).
14. C.D. Schwieters, J. Kuszewski, N. Tjandra, and G.M. Clore. "The Xplor-NIH NMR molecular structure determination package", *J. Magn. Reson.*, **160**, pp. 65–73, (2003).
15. Y. Shen, F. Delaglio, G. Cornilescu, and A. Bax. "TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts", *J. Biomol. NMR*, **44**, pp. 213–223, (2009).
16. Y. Shen, O. Lange, F. Delaglio, P. Rossi, J.M. Aramini, G. Liu, A. Eletsky, Y. Wu, K.K. Singarapu, A. Lemak, A. Ignatchenko, C.H. Arrowsmith, T. Szyperski, G.T. Montelione, D. Baker, D., and A. Bax. "Consistent blind protein structure generation from NMR chemical shift data", *Proc. Natl. Acad. Sci. U.S.A.*, **105**, pp. 4685–4690, (2008).
17. Y. Shen, R. Vernon, D. Baker, and A. Bax. "De novo protein structure generation from incomplete chemical shift assignments", *J. Biomol. NMR*, **43**, pp. 63–78, (2009).
18. T. Szyperski, P. Güntert, G. Otting, and K. Wüthrich. "Determination of scalar coupling constants by inverse Fourier transformation of in-phase multiplets.", *J. Magn. Reson.*, **99**, pp. 552–560, (1992).
19. T.A. Wassenaar, M. van Dijk, N. Loureiro-Ferreira, G. van der Schot, S.J. de Vries, J. van der Zwan, R. Boelens, A. Giachetti, D. Carotenuto, A. Rosato, I. Bertini, T. Herrmann, H.R.A. Jonker, A. Bagaria, V. Jaravine, P. Güntert, H. Schwalbe, W.F.

- Vranken, M. Verlato, S. Dal Pra, M. Mazzucato, and A.M.J.J. Bonvin. “e-NMR: Structural Biology on the Grid”, *Submitted for publication*.
20. W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, and E.D. Laue. “The CCPN data model for NMR spectroscopy: development of a software pipeline”, *Proteins*, **59**, pp. 687–696, (2005).
 21. S.J. de Vries, A.D.J. van Dijk, M. Krzeminski, M. van Dijk, A. Thureau, V. Hsu, T. Wassenaar, and A.M.J.J. Bonvin. “HADDOCK versus HADDOCK: New features and performance of HADDOCK2.0 on the CAPRI targets”, *Proteins: Struc. Funct. & Bioinformatic*, **69**, pp. 726–733, (2007).
 22. S.J. de Vries, M. van Dijk, A.M.J.J. Bonvin. “The HADDOCK server for data-driven biomolecular docking.”, *Nature Protocols*, in press, (2010).
 23. K. Wüthrich. “NMR of Proteins and Nucleic Acids.”, New York, Wiley, (1986).
 24. Y.S. Yung, and M. Zweckstetter. “Mars — robust automatic backbone assignment of proteins”, *J. Biomol. NMR*, **30**, pp. 11–23, (2004).
 25. <http://egee.cesnet.cz/en/JRA1/LB/documentation.php>
 26. <http://glite.web.cern.ch/glite/>
 27. http://glite.web.cern.ch/glite/packages/R3.2/s15_x86_64/deployment/glite-LFC_mysql/glite-LFC_mysql.asp
 28. http://grid-it.cnaf.infn.it/mrepo/ig_s15-x86_64/RPMS.3_2_0/ig-vomscerts-all-latest.noarch.rpm
 29. <http://grid-monitor.pd.infn.it:8008/gstat/summary/GRID/EGEE/>
 30. <http://igrelease.forge.cnaf.infn.it/doku.php>
 31. http://igrelease.forge.cnaf.infn.it/doku.php?id=doc:use_cases:enabling-vo
 32. <http://sbgrid.org/>
 33. <http://twiki.cern.ch/twiki/bin/genpdf/EGEE/InformationSystem>
 34. <http://www.biggrid.nl/>
 35. <http://www.east-nmr.eu/>
 36. <http://www.enmr.eu/>
 37. <http://www.eu-egee.org/>
 38. <http://www.eu-nmr.eu/>
 39. http://www.globus.org/grid_software/security/voms.php
 40. <http://www.gromacs.org/>
 41. <http://www.igtf.net/>
 42. <http://www.italiangrid.org/>
 43. <http://www.ngs.ac.uk/>
 44. <http://www.postgenomicnmr.net/>
 45. <http://www.sagrid.ac.za/>
 46. <http://www.spec.org/>
 47. <http://www.to.infn.it/dgas/index.html>
 48. <http://www3.egee.cesga.es/>
 49. <https://cic.gridops.org>
 50. <https://dgas.cnaf.infn.it/hlrmon/report/charts.php>
 51. <https://e-nmr.cnaf.infn.it/pmwiki.php?n=WP2.1QuickGLiteDeploymentGuide>
 52. <https://e-nmr.cnaf.infn.it/pmwiki.php?n=WP2.2QuickGLiteUserGuide>
 53. <https://edms.cern.ch/document/572489/1>
 54. <https://goc.gridops.org/>
 55. <https://twiki.cern.ch/twiki/bin/view/EGEE/DMEDS>
 56. <https://twiki.cern.ch/twiki/bin/view/EGEE/YAIM>
 57. <https://voms2.cnaf.infn.it:8443/voms/enmr.eu>