

# A success story....

---





## **A worldwide e-Infrastructure for NMR and structural biology**

### **Toward full automation of structure determination by NMR**

---

**10<sup>th</sup> e-Infrastructure Concertation meeting, Brussels March 6-7 2013**



**Alexandre M.J.J. Bonvin**

**Project coordinator**

**Bijvoet Center for Biomolecular Research**

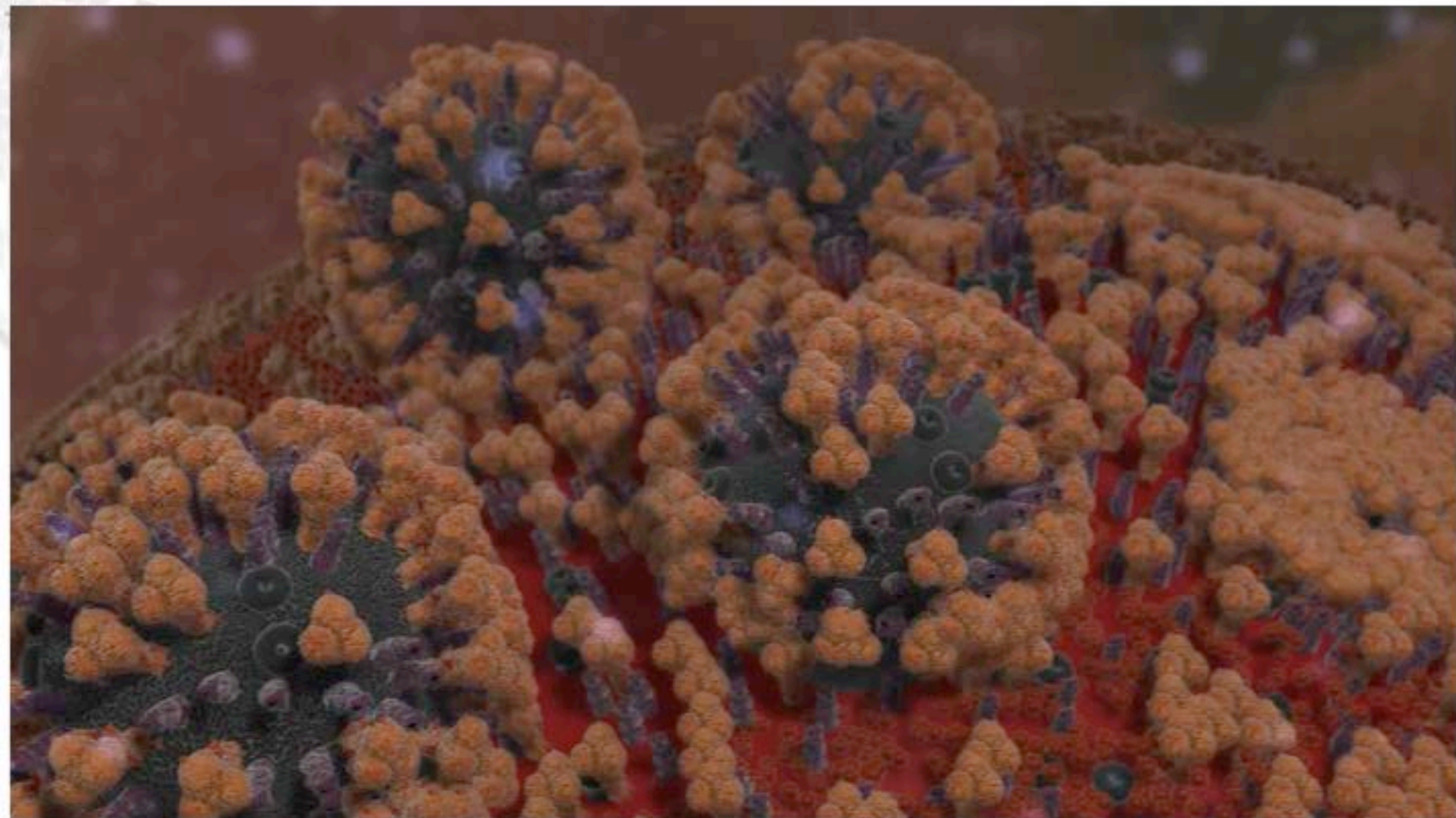
**Faculty of Science, Utrecht University**

**the Netherlands**

**[a.m.j.j.bonvin@uu.nl](mailto:a.m.j.j.bonvin@uu.nl)**



# NMR and structural biology



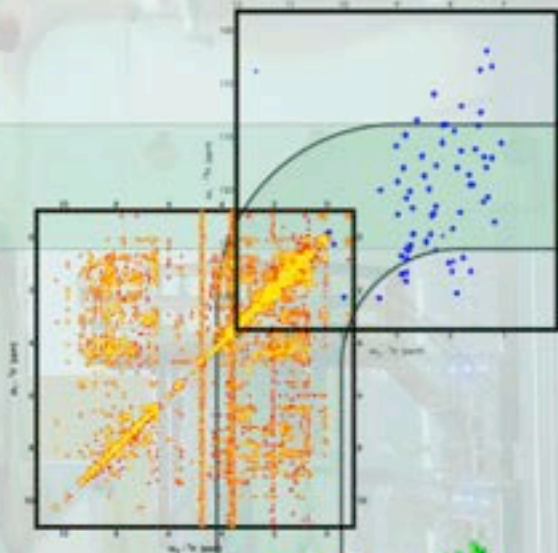


# The NMR structure determination workflow

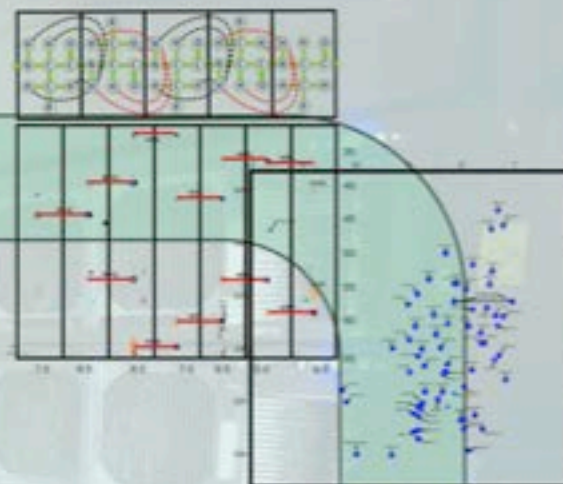
Data acquisition



Data processing

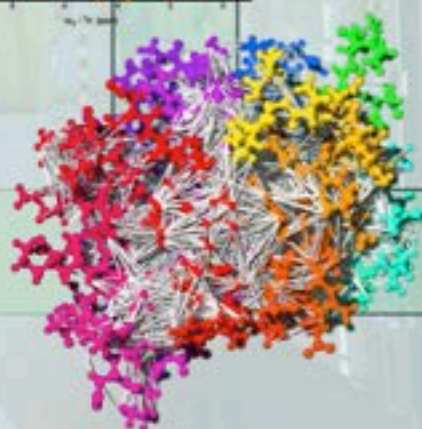


Resonance assignment

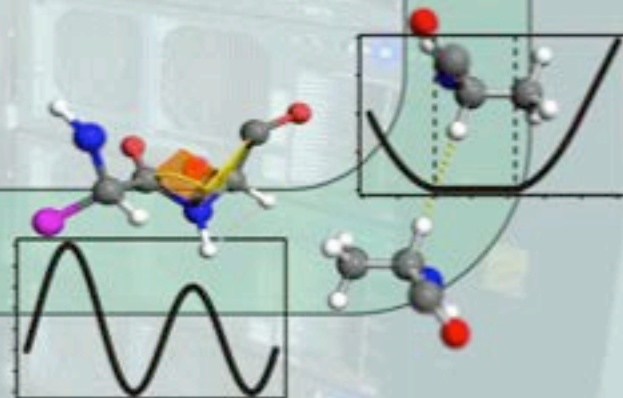


Structure validation

Structure calculations

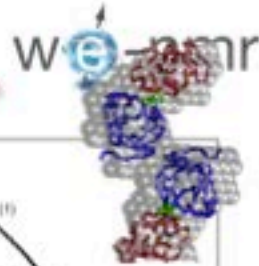


Restraint definitions



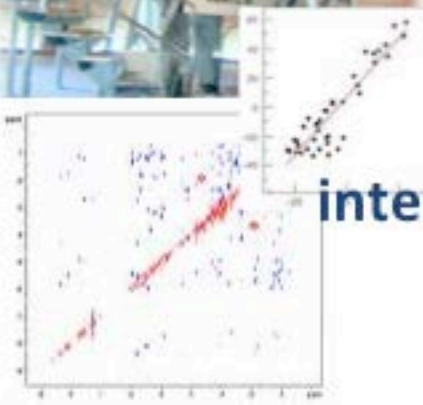
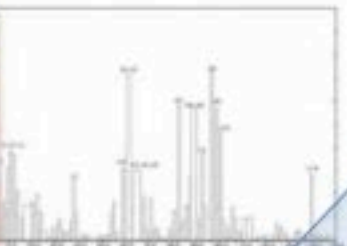
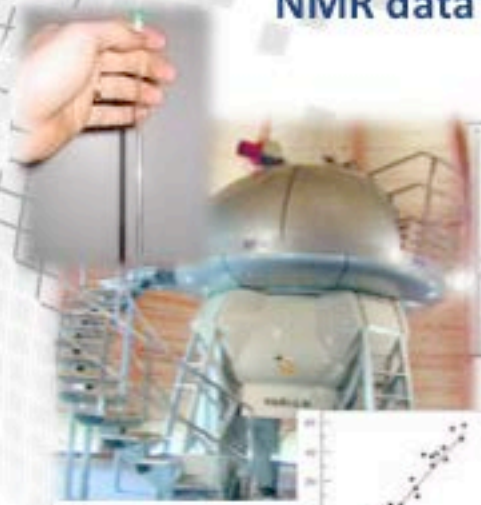


# Exploiting GRID resources in structural biology...



NMR data collection and processing

SAXS data analysis



Data interpretation

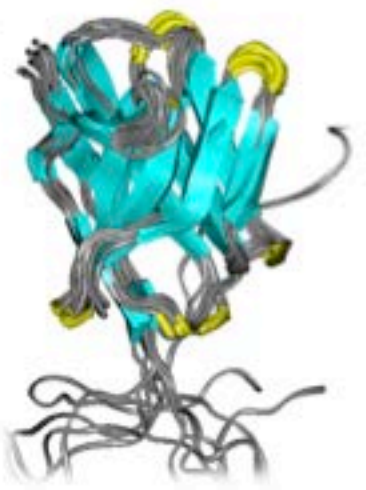
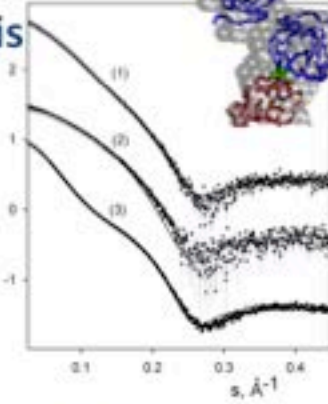
# Number:  
# NAME 1:  
# NAME 2:  
12 2.132 387.17  
14 2.38 401.17  
32 1.849 411.17  
36 1.849 3.143 1.17

Table 2.3: NMR evaluation

Protein group	Score	Rank
Protein group	1.00	1
Protein group	0.99	2
Protein group	0.98	3
Protein group	0.97	4
Protein group	0.96	5
Protein group	0.95	6
Protein group	0.94	7
Protein group	0.93	8
Protein group	0.92	9
Protein group	0.91	10
Protein group	0.90	11
Protein group	0.89	12
Protein group	0.88	13
Protein group	0.87	14
Protein group	0.86	15
Protein group	0.85	16
Protein group	0.84	17
Protein group	0.83	18
Protein group	0.82	19
Protein group	0.81	20
Protein group	0.80	21
Protein group	0.79	22
Protein group	0.78	23
Protein group	0.77	24
Protein group	0.76	25
Protein group	0.75	26
Protein group	0.74	27
Protein group	0.73	28
Protein group	0.72	29
Protein group	0.71	30
Protein group	0.70	31
Protein group	0.69	32
Protein group	0.68	33
Protein group	0.67	34
Protein group	0.66	35
Protein group	0.65	36
Protein group	0.64	37
Protein group	0.63	38
Protein group	0.62	39
Protein group	0.61	40
Protein group	0.60	41
Protein group	0.59	42
Protein group	0.58	43
Protein group	0.57	44
Protein group	0.56	45
Protein group	0.55	46
Protein group	0.54	47
Protein group	0.53	48
Protein group	0.52	49
Protein group	0.51	50
Protein group	0.50	51
Protein group	0.49	52
Protein group	0.48	53
Protein group	0.47	54
Protein group	0.46	55
Protein group	0.45	56
Protein group	0.44	57
Protein group	0.43	58
Protein group	0.42	59
Protein group	0.41	60
Protein group	0.40	61
Protein group	0.39	62
Protein group	0.38	63
Protein group	0.37	64
Protein group	0.36	65
Protein group	0.35	66
Protein group	0.34	67
Protein group	0.33	68
Protein group	0.32	69
Protein group	0.31	70
Protein group	0.30	71
Protein group	0.29	72
Protein group	0.28	73
Protein group	0.27	74
Protein group	0.26	75
Protein group	0.25	76
Protein group	0.24	77
Protein group	0.23	78
Protein group	0.22	79
Protein group	0.21	80
Protein group	0.20	81
Protein group	0.19	82
Protein group	0.18	83
Protein group	0.17	84
Protein group	0.16	85
Protein group	0.15	86
Protein group	0.14	87
Protein group	0.13	88
Protein group	0.12	89
Protein group	0.11	90
Protein group	0.10	91
Protein group	0.09	92
Protein group	0.08	93
Protein group	0.07	94
Protein group	0.06	95
Protein group	0.05	96
Protein group	0.04	97
Protein group	0.03	98
Protein group	0.02	99
Protein group	0.01	100



Computations



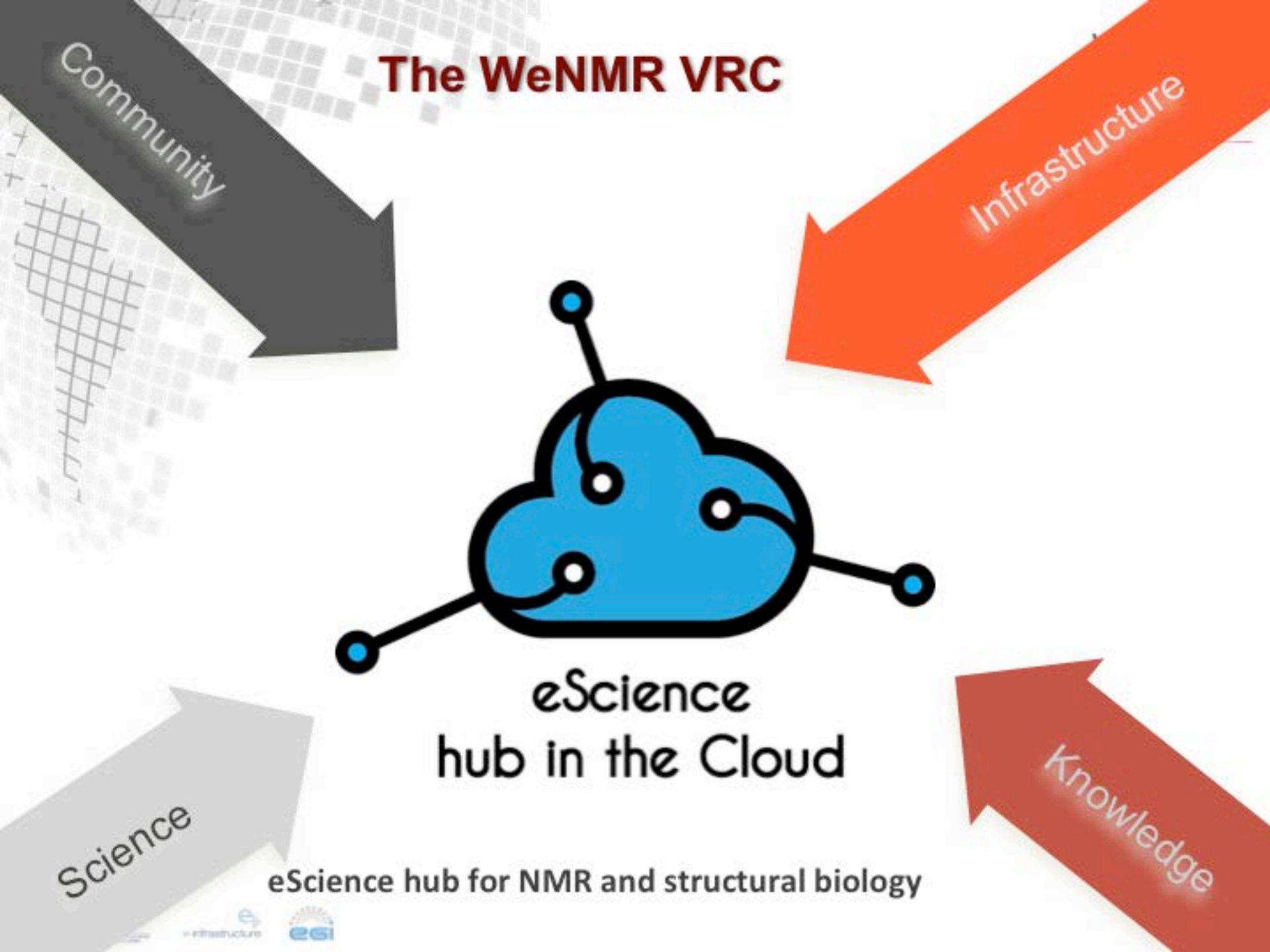
Structure, dynamics & interactions

→ impact on research and health:

- origin of disease
- design of new experiments
- drug design...



# The WeNMR VRC

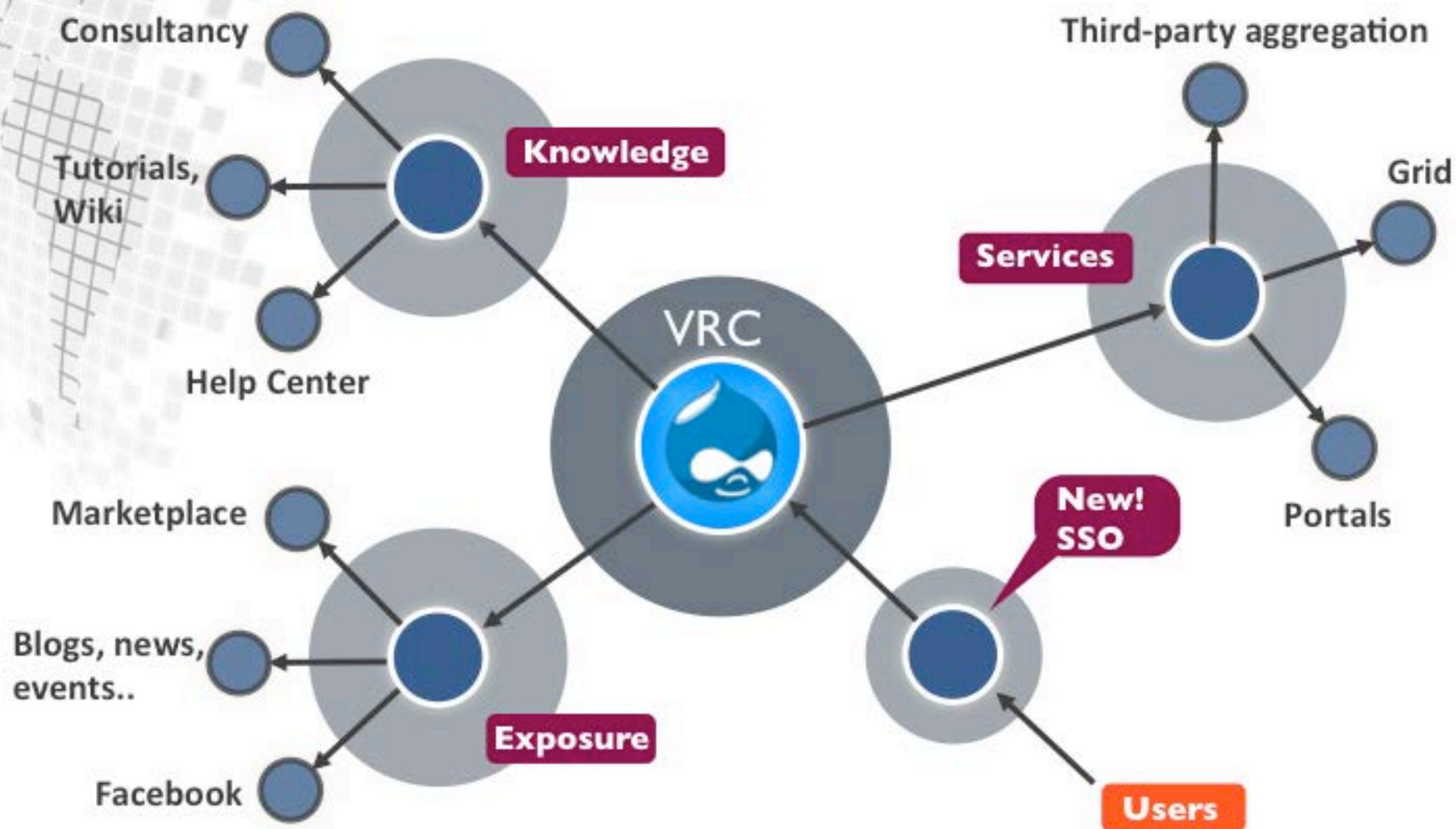


eScience  
hub in the Cloud

eScience hub for NMR and structural biology



# The WeNMR VRC





# The WeNMR services portfolio



The image displays a collection of web portals for various bioinformatics and structural biology services, all part of the WeNMR portfolio. The portals are arranged in a collage-like fashion, overlapping each other. Each portal has a blue header with the service name and a 'WeNMR GRID-enabled web portal' tagline. The services shown include:

- HADDOCK**: A web portal for protein-protein docking.
- Xplor-NIH**: A web portal for NMR data processing and analysis.
- AMBER**: A web portal for molecular dynamics simulations.
- CYANA**: A web portal for NMR data processing and analysis.
- CS-ROSETTA**: A web portal for protein structure prediction.
- MAXOCC**: A web portal for NMR data processing and analysis.
- GROMACS**: A web portal for molecular dynamics simulations.
- CCPN format**: A web portal for NMR data processing and analysis.
- 3DDART**: A web portal for NMR data processing and analysis.
- MARS**: A web portal for NMR data processing and analysis.
- Antechamber**: A web portal for NMR data processing and analysis.
- ASDP**: A web portal for NMR data processing and analysis.
- SHIFTX2**: A web portal for NMR data processing and analysis.

Each portal typically features a navigation menu, a main content area with text and images, and a footer with contact information. The overall design is professional and scientific, with a focus on providing accessible web-based interfaces for complex computational tasks.





## WeNMR platform (February 2013)

- Largest VO in the life sciences
- > 500 registered users (29% outside the EU)
- ~ 45 000 CPU cores
- > 330 CPU years over the last 12 months
- > 1.25 million jobs over the last 12 months
- User-friendly access to Grid via web portals



[www.wenmr.eu](http://www.wenmr.eu)

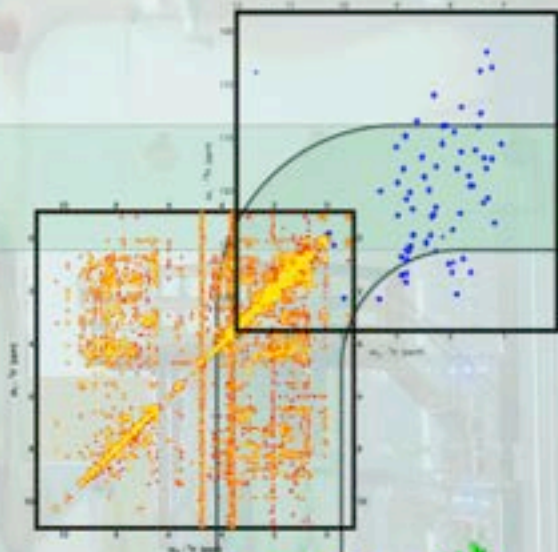


# The NMR structure determination workflow

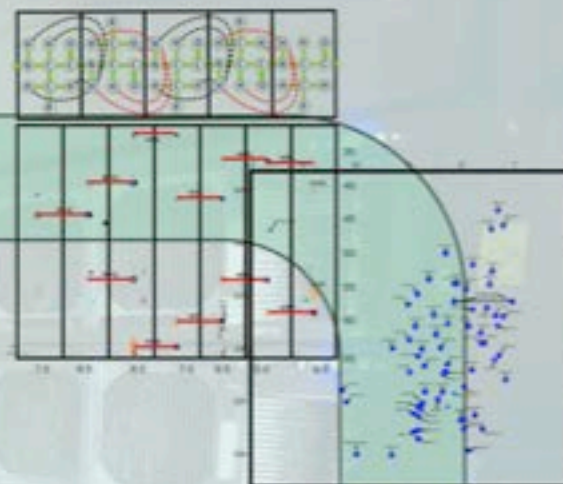
Data acquisition



Data processing

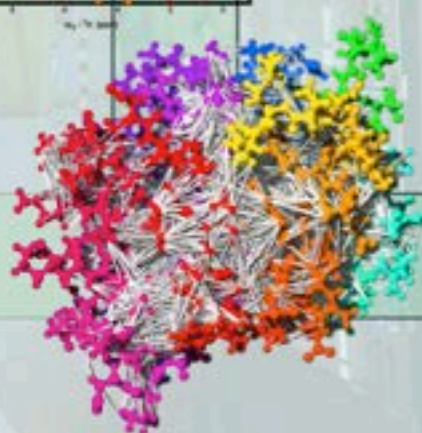


Resonance assignment

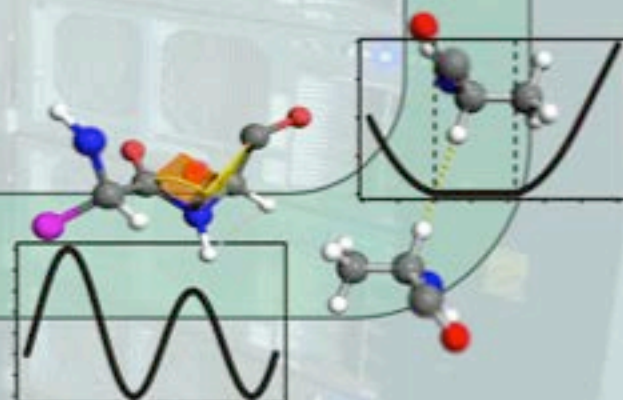


Structure validation

Structure calculations



Restraint definitions



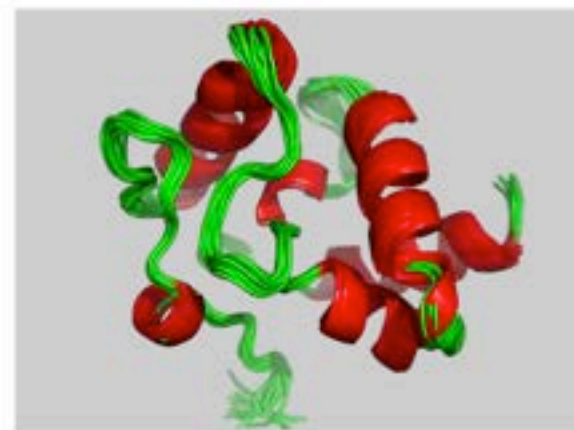


*Nature Methods* **6**, 625 - 626 (2009)  
doi:10.1038/nmeth0909-625

## CASD-NMR: critical assessment of automated structure determination by NMR

Antonio Rosato<sup>1,2</sup>, Anurag Bagaria<sup>3,4</sup>, David Baker<sup>5</sup>, Benjamin Bardiaux<sup>6</sup>, Andrea Cavalli<sup>7</sup>, Jurgen F Doreleijers<sup>8</sup>, Andrea Giachetti<sup>1</sup>, Paul Guerry<sup>9</sup>, Peter Güntert<sup>3,4</sup>, Torsten Herrmann<sup>9</sup>, Yuanpeng J Huang<sup>10</sup>, Hendrik R A Jonker<sup>4,11</sup>, Binchen Mao<sup>10</sup>, Thérèse E Malliavin<sup>6</sup>, Gaetano T Montelione<sup>10</sup>, Michael Nilges<sup>6</sup>, Srivatsan Raman<sup>5</sup>, Gijs van der Schot<sup>12</sup>, Wim F Vranken<sup>13</sup>, Geerten W Vuister<sup>8</sup> & Alexandre M J J Bonvin<sup>12</sup>

**nature** | **methods**  
Techniques for life scientists and chemists



- **10 blind targets** (structures to be determined by fully automated methods)
- **15 teams worldwide** involved with their respective developer teams
- Assessment of results by independent team
- **Impact: WeNMR & EU as drivers of innovation, leading to to new software developments, increased trust in automated methods and improved services for end users**



## Structure

## Ways &amp; Means

Structure 20, 227–236, 2012

## Blind Testing of Routine, Fully Automated Determination of Protein Structures from NMR Data

Antonio Rosato,<sup>1,2,4</sup> James M. Aramini,<sup>3</sup> Cheryl Arrowsmith,<sup>4</sup> Anurag Bagaria,<sup>5,6</sup> David Baker,<sup>8</sup> Andrea Cavalli,<sup>9</sup> Jürgen F. Doreleijers,<sup>10</sup> Alexander Eletsky,<sup>11</sup> Andrea Giachetti,<sup>1</sup> Paul Guerry,<sup>12</sup> Aleksandras Gutmanas,<sup>4</sup> Peter Yunfen He,<sup>11</sup> Torsten Herrmann,<sup>12</sup> Yuanpeng J. Huang,<sup>3</sup> Victor Jaravine,<sup>5,6</sup> Hendrik R.A. Jonker,<sup>6,7</sup> Michael A. J. Oliver F. Lange,<sup>8</sup> Gaohua Liu,<sup>3</sup> Thérèse E. Malliaris,<sup>14</sup> Rajeswari Mari,<sup>3</sup> Binchen Mao,<sup>3</sup> Gaetano T. Montellon,<sup>15</sup> Michael Nilges,<sup>14</sup> Paolo Rossi,<sup>3</sup> Gijs van der Schot,<sup>15</sup> Harald Schwalbe,<sup>6,7</sup> Thomas A. Szyperski,<sup>11</sup> Michele Ver Robert Vernon,<sup>8</sup> Wim F. Vranken,<sup>16,17</sup> Sjoerd de Vries,<sup>15,18</sup> Geerten W. Vuister,<sup>10,18</sup> Bin Wu,<sup>4</sup> Yunhuang Yang, and Alexandre M.J.J. Bonvin<sup>15</sup>

<sup>1</sup>Magnetic Resonance Center<sup>2</sup>Department of Chemistry

University of Florence, 50019 Sesto Fiorentino, Italy

<sup>3</sup>Center for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry, and Northeast Structural Consortium (NESG), Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA<sup>4</sup>Ontario Cancer Institute, NESG, University of Toronto, Toronto, Ontario M5G 2M9, Canada<sup>5</sup>Institute of Biophysical Chemistry and Frankfurt Institute for Advanced<sup>6</sup>Center for Biomolecular Magnetic Resonance<sup>7</sup>Institute of Organic Chemistry and Chemical Biology

Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

<sup>8</sup>Department of Biochemistry, University of Washington, Seattle, WA 98<sup>9</sup>Department of Chemistry, University of Cambridge, Cambridge CB2 1<sup>10</sup>Protein Biophysics, Institute of Molecules and Materials, Radboud Un<sup>11</sup>Department of Chemistry, NESG, University at Buffalo, The State Uni<sup>12</sup>Centre de RMN à très Hauts Champs, Institut des Sciences Analytiq

69100 Villeurbanne, France

<sup>13</sup>Department of Chemistry and Biochemistry, and NESG, Miami Unive

research advances

## Protein Structure Initiative

April 2012 technical highlight

## Blind faith

SBKB [doi:10.1038/sbkb.2011.73]

A rigorous assessment of automated NMR structure determination demonstrates the reliability of the method.

Protein structure determination by NMR spectroscopy requires complete chemical shift assignments and manual analysis of thousands of atomic interactions known as nuclear Overhauser effects (NOEs). Recently developed algorithms use these data to calculate structures in automated, iterative NOE assignment protocols. Additional methods have emerged that rely on the intrinsic structural information encoded in the chemical shifts supplemented with NOE data.

Rosato and colleagues (PSI NESG) present the results of CASD-NMR, a blind assessment of how the available automated methods compare with the traditional manual approach. The authors stringently evaluated the results for ten structure determinations of monomeric proteins ranging in size from 60 to 150 amino acids. They assessed three groups of automated methods using either NOE or chemical shift data, or a combination of both. The accuracy and convergence of the automatically calculated structures was compared to manually solved reference structures using the root-mean-square deviation (RMSD) of structured backbone regions.

Setting a threshold for accuracy at an RMSD of 2 Å from



European Grid Infrastructure

towards a sustainable infrastructure

COMMUNITY • INFRASTRUCTURE • SERVICES • HOW DO I...? • CASE STUDIES • NEWS &amp; MEDIA • ABOUT

Newsfeed

Newsletters

Presentations

Publications

Director's letters

Press

Videos

HOME • NEWS &amp; MEDIA • NEWSFEED • WANT TO KNOW WHAT A MOLECULE LOOKS LIKE? CHALLENGE A COMPUTER!

## Want to know what a molecule looks like? Challenge a computer!

The first results have been published from a public challenge to automate the process of determining the shape and structure of a molecule from NMR data. Three of the teams who took up the challenge have now posted their work to the grid.

NMR (Nuclear Magnetic Resonance) is a technique used by many researchers to investigate how the constituents of a molecule interact and the shapes they form. It does this by exploiting a technique where one can 'listen' to atoms if one puts them into a very strong magnetic field. However, the music they play is complex and it requires a lot of work to interpret the NMR raw data obtained from a molecule before one can figure out its exact shape.

13 April 2012  
Neenan O'Neill



# Acknowledgments

## The team

Alexandre M.J.J. Bonvin  
Andrea Giachetti  
Antonio Rosato  
Anurag Bagaria  
Christophe Schmitz  
Eric Frizziero  
Gijs van der Schot  
Harald Schwalbe  
Hendrik R. A. Jonker

Ivano Bertini  
Johan van der Zwan  
Lucio Ferella  
Marc van Dijk  
Marco Verlato  
Mikael Trellet  
Mirco Mazzucato  
Nuno Loureiro-Ferreira  
Peter Güntert

Rolf Boelens  
Sjoerd J. de Vries  
Stefano Dal Pra  
Torsten Herrmann  
Tsjerk A. Wassenaar  
Victor Jaravine  
Wim F. Vranken  
Danny Hsu  
Simon Lin

## The partners

Universiteit Utrecht



Utrecht University, Bijvoet Center for Biomolecular Research, NL



Johann Wolfgang Goethe Universität Frankfurt a.M., Center for Biomolecular Magnetic Resonance DE



University of Florence, Magnetic Resonance Center, IT

Istituto Nazionale di Fisica Nucleare, Padova, IT



Raboud University, Nijmegen, NL



University of Cambridge UK

EMBL



European Molecular Biology Laboratory, Hamburg, DE

Spronk NMR Consultancy, LT

Academia Sinica



**Contract n°:** RI-261572  
**Project type:** CP-CSA





