DEISA

Distributed European Infrastructure for Supercomputing Applications

Wolfgang Gentzsch
gentzsch@rzg.mpg.de

Cracow Grid Workshop
October 13, 2008

www.deisa.eu

RI-222919

DEISA

Distributed European Infrastructure for Supercomputing **Applications**

Wolfgang Gentzsch
gentzsch@rzg.mpg.de

Cracow Grid Workshop
October 13, 2008

www.deisa.eu

HPC Users Requirements

- **Remote, Standard, Easy Access**

  HPC users: conservative, no interest in complicated middleware stack
- **Global Login**

  Personal Login in each system, I.e. single "European" username, uniform access
- **Comfortable Data Access**

  Global, fast and comfortable access to their data, across the centres
- **Common Production Environment**

  No need for an identical but for an equivalent HPC software stack
- **Global Help Desk**

  One central point of contact and as local as possible
- **Application Support**

  Need help in scalability, performance and adaptation to different architectures

Sustainability? Carefully listen to the top user value propositions and go build and 'sell' it!

Objectives of the European DEISA1 Project

Distributed European Infrastructure for Supercomputing Applications

- **Enabling Europe's terascale science by integrating Europe's most powerful HPC systems**

  DEISA is a European Supercomputing Service built on top of existing national HPC services This Service is based on the deployment and operation of a persistent, production quality, distributed supercomputing environment with continental scope

- **Only Criterion for Success: Enabling scientific discovery across a broad spectrum of science and technology**

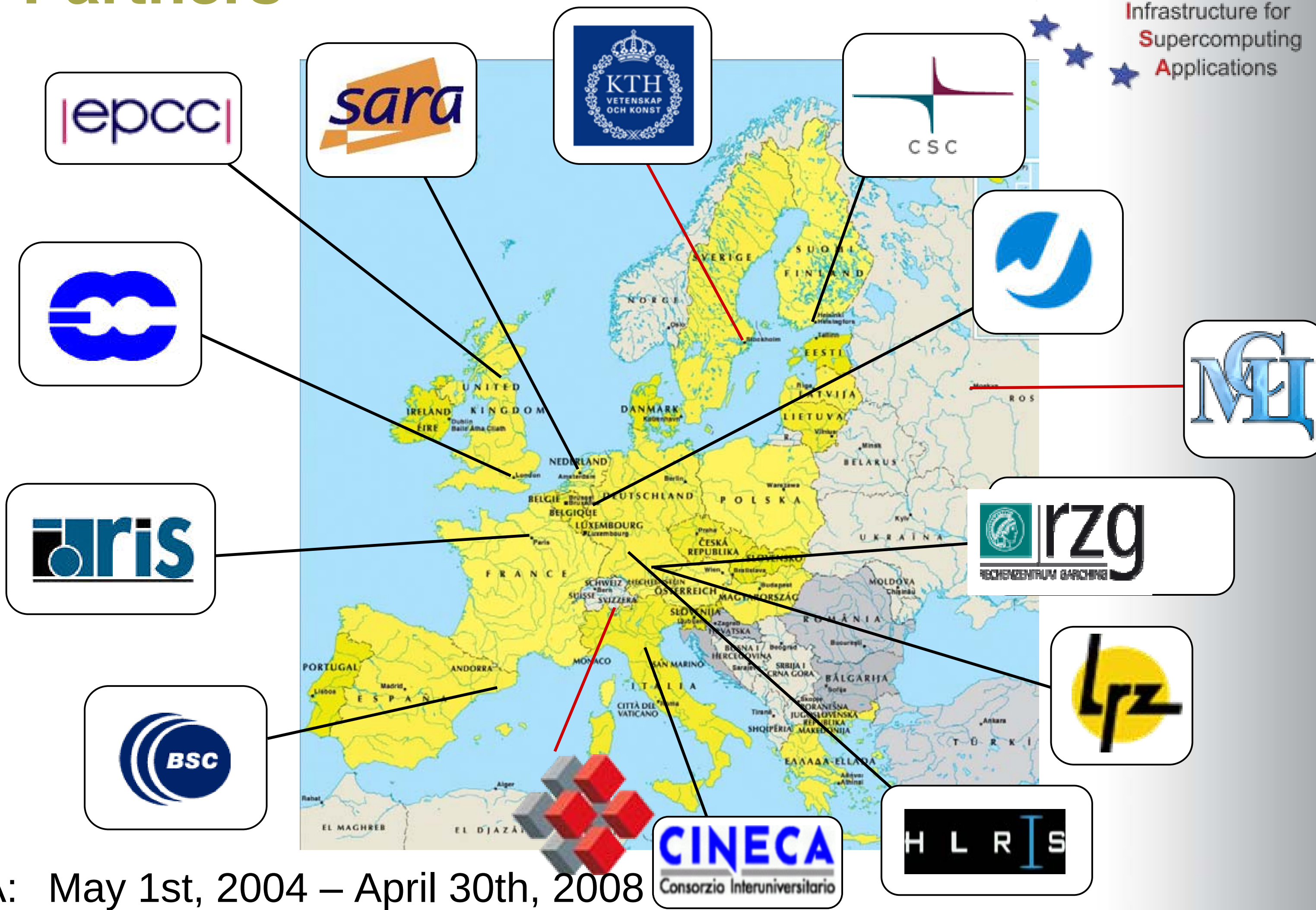
  The integration of national facilities and services, together with innovative operational models, is expected to add substantial value to existing infrastructures

Objectives
of the European DEISA2 Project

- Enhancing the existing distributed European HPC (DEISA1) environment towards a turnkey operational infrastructure
- Advancing computational sciences in Europe by supporting user communities and extreme computing projects
- Enhancing service provision by offering a manageable variety of options of interaction with computational resources
- Integration of T1 and T0 centers
- The Petascale T0 Systems need a transparent access from and into the national data repositories
- Bridging worldwide HPC and Grid projects

Partners

Distributed
European
Infrastructure for
Supercomputing
Applications



DEISA: May 1st, 2004 – April 30th, 2008

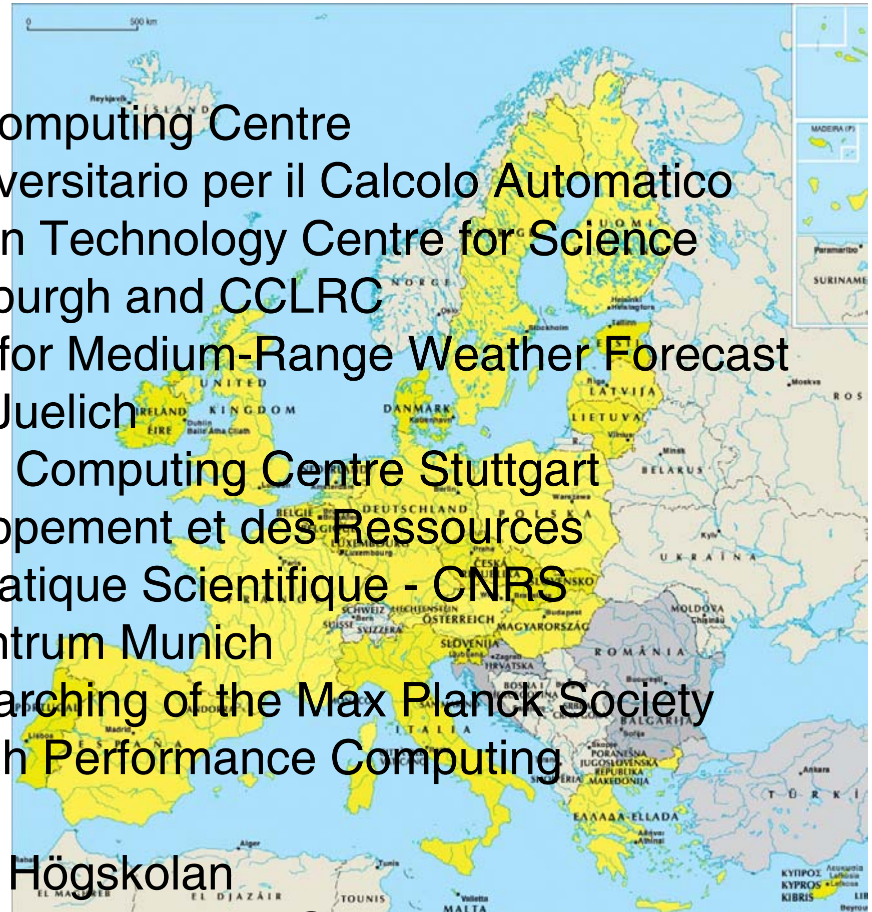
Three new partners joined June 2005 (eDEISA)

DEISA2: May 1st, 2008 – April 30th, 2011

RI-222919

Distributed European Infrastructure for Supercomputing Applications

Partners



BSC	Barcelona Supercomputing Centre	Spain
CINECA	Consortio Interuniversitario per il Calcolo Automatico	Italy
CSC	Finnish Information Technology Centre for Science	Finland
EPCC	University of Edinburgh and CCLRC	UK
ECMWF	European Centre for Medium-Range Weather Forecast	UK (int)
FZJ	Research Centre Juelich	Germany
HLRS	High Performance Computing Centre Stuttgart	Germany
IDRIS	Institut du Développement et des Ressources en Informatique Scientifique - CNRS	France
LRZ	Leibniz Rechenzentrum Munich	Germany
RZG	Rechenzentrum Garching of the Max Planck Society	Germany
SARA	Dutch National High Performance Computing	Netherlands
KTH	Kungliga Tekniska Högskolan	Sweden
CSCS	Swiss National Supercomputing Centre	Switzerland
JSCC	Joint Supercomputer Center of the Russian Academy of Sciences	Russia

How is DEISA Enhancing HPC Services in Europe?

Distributed European Infrastructure for Supercomputing Applications

- Running large parallel applications in individual sites, orchestrating the global workload, or by job migration services
- Enabling workflow applications with UNICORE (complex applications that are pipelined over several computing platforms), enabling coupled multi-physics Grid applications
- Providing a global data management service whose main objectives are:
  - Integrating distributed data with distributed computing platforms
  - Enabling efficient, high performance access to remote datasets with Global File System and striped GridFTP.
  - Integrating hierarchical storage mgmt and databases in the Grid.
- Deploying portals to hide complex environments to new users, and to interoperate with other existing grid infrastructures.

DEISA Services

Multiple ways to access	Workflow managem.	Common production environm.	*Presentation layer*
Single monitor system	Job rerouting	Co-reservation and co-allocation	*Job manag. layer and monitor.*
Data staging tools	Data transfer tools	WAN shared filesystem	*Data manag. layer*
Unified AAA	DEISA Sites	Network connectivity	*Network and AAA layers*

Multiple ways to access

Workflow managem.

Common production environm.

DEISA
a service-oriented production (RAS) e-Infrastructure

Si
mo
sy

tion
o-
ion

D
sta
to

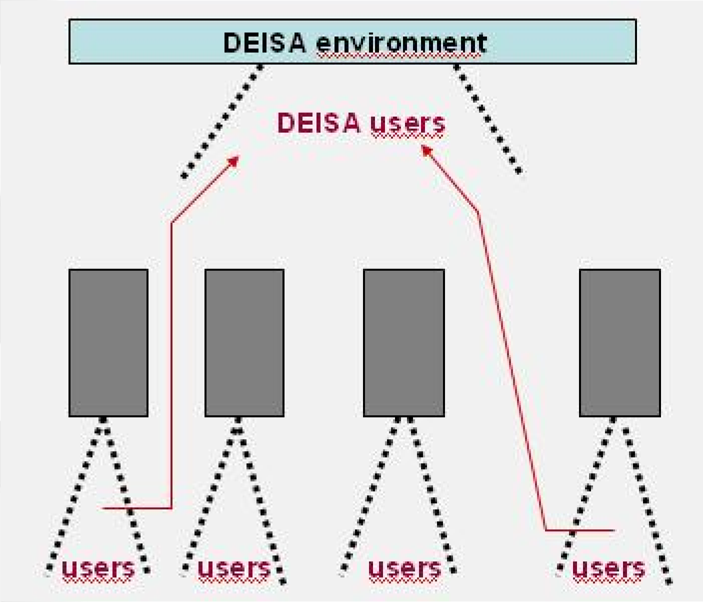
N
ed
stem

Unified AAA

DEISA Sites

Network connectivity

Enabling Science: DEISA users communities



National users communities have accounts on a given site and do not « naturally » see the whole DEISA environment.

Promotion of DEISA users is done via the Extreme Computing Initiative.

Europ. call for proposals for grand challenge simulations every year in May since 2005.

About 50 grand challenge projects supported each year since 2005.

Full information about Extreme Computing projects and reports from terminated projects can be found on the DEISA Web server: www.deisa.eu

The Extreme Computing Initiative is the current DEISA service provisioning model.

DEISA Core Infrastructure and Services

Distributed European Infrastructure for Supercomputing Applications

- Dedicated high speed network
- Common AAA
- Single sign on
  - Accounting/budgeting
- Global data management
- High performance remote I/O and data sharing with global file systems
  - High performance transfers of large data sets
- User Operational infrastructure
  - Distributed Common Production Environment (DCPE)
  - Job management service
  - Common user support and help desk
- System Operational infrastructure
  - Common monitoring and information systems
  - Common system operation
- Global Application Support

DEISA Extreme Computing Initiative DECI

Distributed European Infrastructure for Supercomputing Applications

Application Enabling and Optimizations

- Scaling of parallel programs for efficient usage on thousands of processor-cores challenging, necessary task for state-of-the-art supercomputers

- Support for data intensive applications

- Design, deployment and optimization of workflow applications to chain several compute tasks (simulation, pre- and post-processing steps)

- Coupled applications important e.g. for climate system modelling with separate components for ocean, atmosphere, sea ice, soil, atmospheric chemistry and aerosols

- Adaptation to the DEISA infrastructure
- Optmizations, also architecture dependent optimizations
- Determination of best suited architecture

DEISA Extreme Computing Initiative

Distributed European Infrastructure for Supercomputing Applications

Annual Calls for Proposals for challenging supercomputing projects from all areas of science

DECI call 2005

> 50 proposals, 12 European countries involved, co-investigators from US
> 30 mio cpu-h requested
28 projects selected, 12 mio cpu-h awarded

DECI call 2006

> 40 proposals, 12 European countries involved, co-investigators from N + S America, Asia (US, CA, AR, ISRAEL)
~ 30 mio cpu-h requested
23 projects selected, 12 mio cpu-h awarded

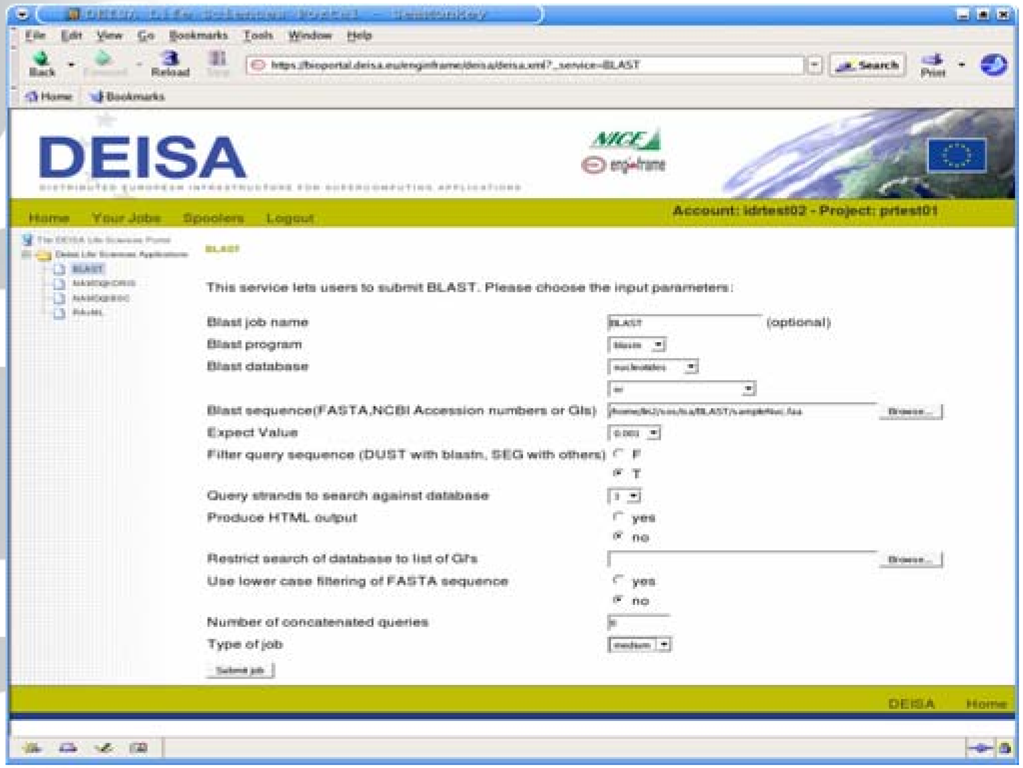
DECI call 2007

> 60 proposals, 14 European countries involved, co-investigators from N + S America, Asia, Australia (US, CA, BR, AR, ISRAEL, AUS)
> 70 mio cpu-h requested
45 projects selected, ~ 30 mio cpu-h awarded

DECI call 2008

- 66 proposals, 14 European countries, USA (7), China, Japan involved
- 134 mio cpu-h requested

One Example: Joint Research Activity "Life Sciences"



DEISA Life Science Portal

Promoting parallel apps in the life science community

Running big simulations on DEISA infrastructure that couldn't be done locally

Providing ease of access to resources

Application support for life science portal

Distributed European Infrastructure for Supercomputing Applications

The Life Sciences Portal

- Objective: reaching new users communities that have already structured their application strategies around small discipline-oriented grid infrastructures with discipline specific tools.
- Strategy: connect supercomputer environments as « backend » resources to existing discipline-oriented eInfrastructures.
- 2006-2007, DEISA has deployed a portal to a European HPC service for Bio-Informatics / Life Sciences with a need for HPC.
- Critical domain applications are ported to best suited supercomputer(s) of the DEISA environment.
- Shared distributed data repositories are hosted by GPFS services.

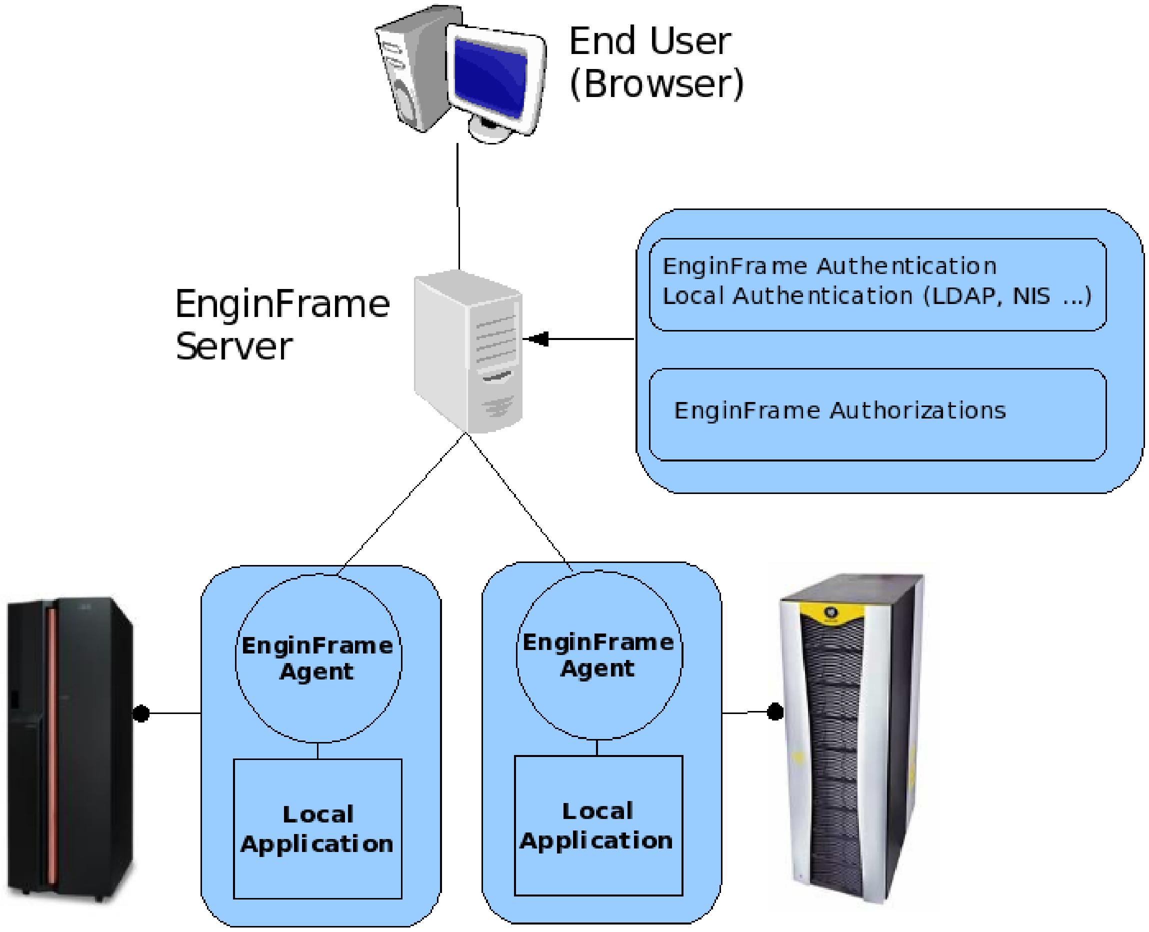
RI-222919

Distributed European Infrastructure for Supercomputing Applications

Portal Deployment

Site	Application	Computing Nodes Architecture	Submission Host Architecture	Scheduler	EF Agent / SSH Plugin
BSC	NAMD	IBM Power PC Linux	i586 Suse Linux 10	SLURM/MOAB	SSH
IDRIS	BLAST	IBM Power6 AIX	IBM Power6 AIX	LoadLeveler	Agent
LRZ	RAxML	SGI Altix Linux	Intel Itanium 2 Suse Linux 9	PBSPRO	Agent
CINECA	BLAST	IBM Power5 AIX	IBM Power 5 AIX 5.3	LoadLeveler	SSH
FZJ	BLAST	IBM Power4+ AIX	IBM Power 4+ AIX 5.3	LoadLeveler	SSH
RZG	BLAST	IBM Power4 AIX	IBM Power 4 AIX 5.3	LoadLeveler	Agent
EPCC	BLAST	IBM Power5 AIX	Power 5 AIX 5.3	LoadLeveler	SSH
CSC	NAMD	CRAY XT4 Opteron Unicos	Opteron Suse Linux 9	PBSPRO	Agent
SARA	NAMD	IBM Power 5 Linux	IBM Power 5 Suse Linux 9	LoadLeveler	Agent

DEISA Life Science Portal based on NICE EnginFrame



RI-222919

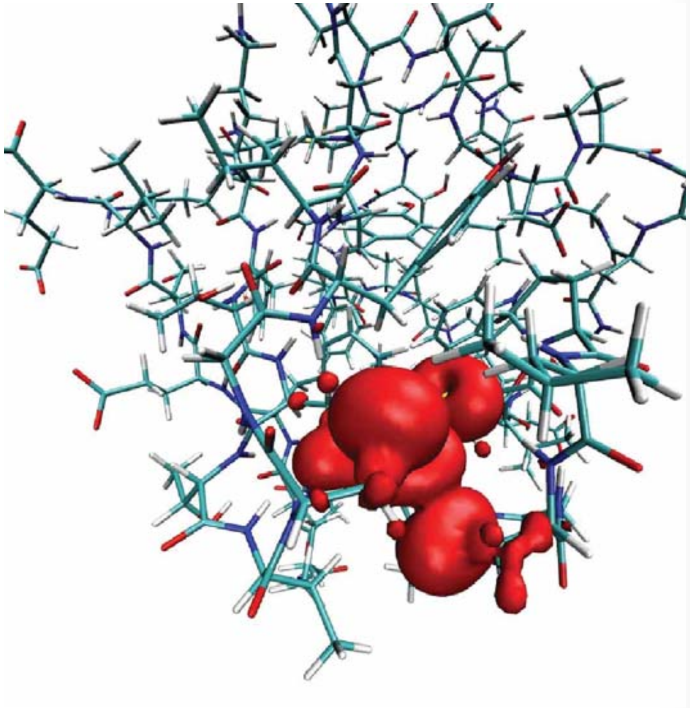
Biological electron transfer simulated

A RESEARCHER GROUP HAS USED FIRST-PRINCIPLES TECHNIQUES TO INVESTIGATE ELECTRON TRANSFER IN THE CONTEXT OF IRON- AND COPPER-BASED PROTEINS.

FOR THE FIRST TIME AB INITIO INVESTIGATION OF ENTIRE ELECTRON-TRANSFER PROTEINS IN AQUEOUS SOLUTION WAS PERFORMED.

Isosurface enclosing 95 per cent of the total spin density (difference of the up and down electron densities) in the oxidized form of the rubredoxin from Clostridium pasteurianum.

All of the atoms were explicitly taken into account at ab initio (DFT) level.



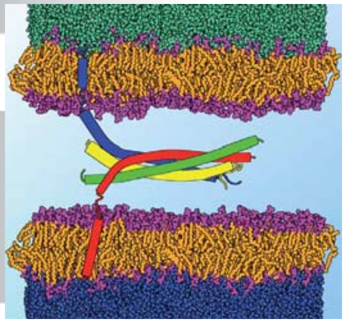
Membranes under tension

Distributed European Infrastructure for Supercomputing Applications

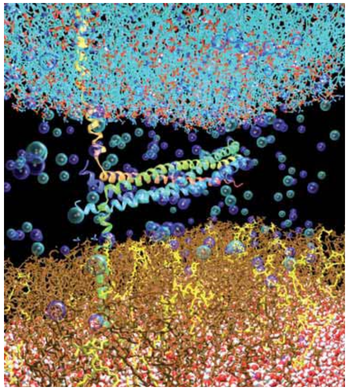
MODELLING OF A PROTEIN COMPLEX THAT ACTS AS A CATALYST IN THE FUSION OF TWO CELL MEMBRANES.

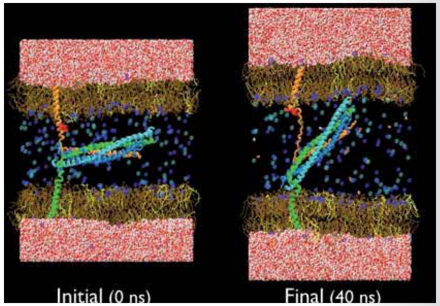
UNDERSTANDING THE GUIDING PRINCIPLES OF MEMBRANE FUSION WILL OPEN NEW OPPORTUNITIES FOR PHARMACEUTICAL DEVELOPMENT.

Perspective view on the membrane embedded SNARE system, highlighting the counterions Na+ and Cl- as bubbles. They are required to neutralize the high charge density of the mixed POPC (brown) / POPS (yellow) membrane.



Schematic view of the SNARE complex embedded in two membranes. The upper membrane represents the vesicle, the lower membrane is the target with which fusion occurs.



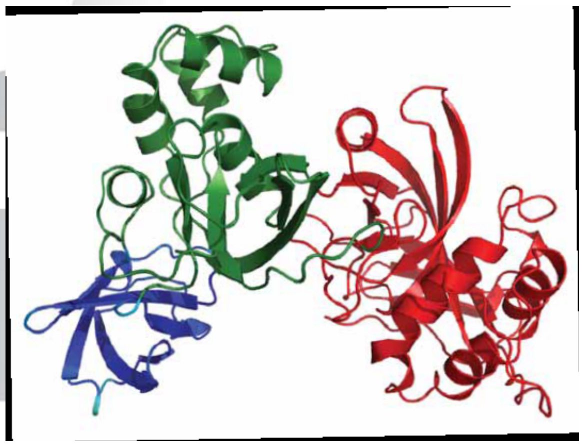


Initial and final structures of simulations on the membrane embedded SNARE complex. Tryptophanes 89 and 90 of synaptobrevin (orange) are highlighted in red. The composition of the membranes is visible with neutral POPC in brown and charged POPS in yellow.

RI-222919

Molecular simulation to accelerate vaccine development

A LARGE THREE-YEAR PROJECT INVOLVING PHARMACEUTICAL AND BIOINFORMATICS COMPANIES AND UNIVERSITIES HAS EXPLOITED THE DEISA RESOURCES TO RENDER VACCINE DEVELOPMENT SIMPLER, FASTER AND THEREFORE LESS EXPENSIVE FOR PHARMACEUTICAL COMPANIES.



To study protein dynamics, classical molecular dynamics simulation in combination with empirical biomolecular force fields have been used.

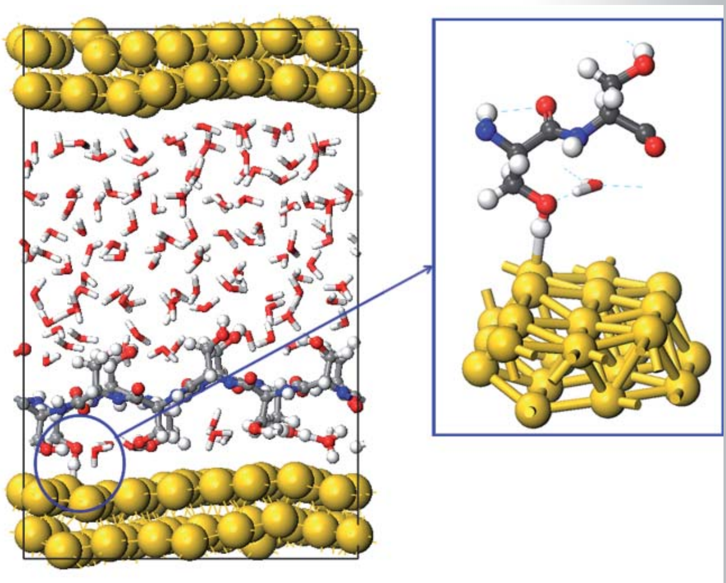
High throughput cloning and expression of large sets of genomic ORFs has become a preferred industrial strategy for genome-wide searches of new vaccine candidates. For invasive infections in particular, the aim is to find proteins eliciting antibodies capable of binding to the bacterial cell surface and, through interaction with the complement system, effectively kill the bacteria.

Protein–surface interactions mediated by water

THE MOLECULAR RECOGNITION CAPABILITY OF PROTEINS CAN BE SPECIFICALLY ORIENTED TOWARD INORGANIC SURFACES. THIS PROJECT AIMS TO ELUCIDATE THE MECHANISMS OF INTERACTION BETWEEN A SURFACE AND A PROTEIN, INCLUDING THE ROLE PLAYED BY WATER.

Snapshot from ab initio molecular dynamics of protein–gold–water system.

The atoms represented (left) are repeated by 3D periodic boundary conditions, creating an infinite gold slab and an infinite poly-Serine ß-sheet. The right highlights the instantaneous interaction between the Serine side-chain and the gold surface, which takes place via the hydrogen atom of the hydroxyl group.

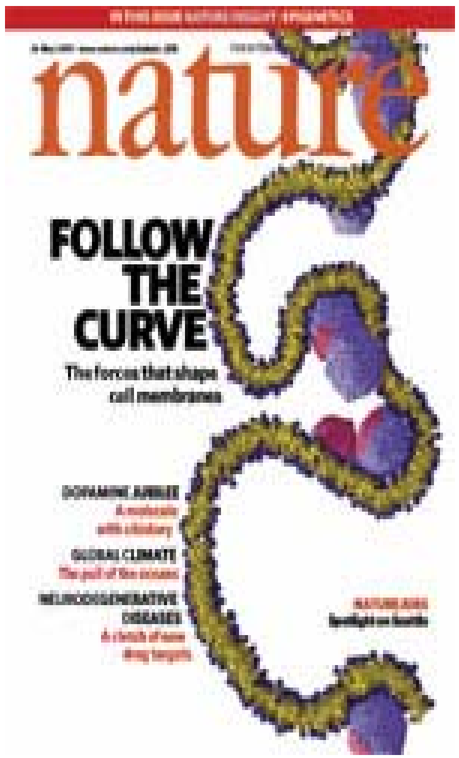


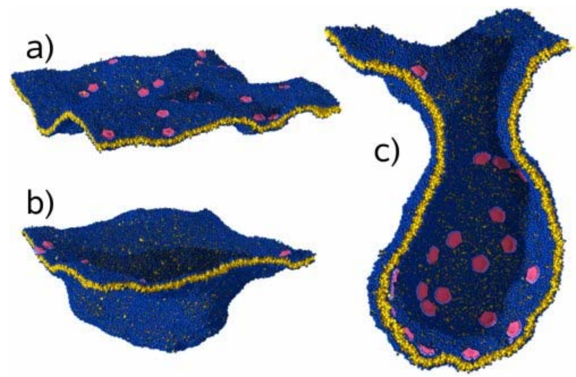
The dynamics of gold atoms are important for explaining the interaction with the protein and with water.

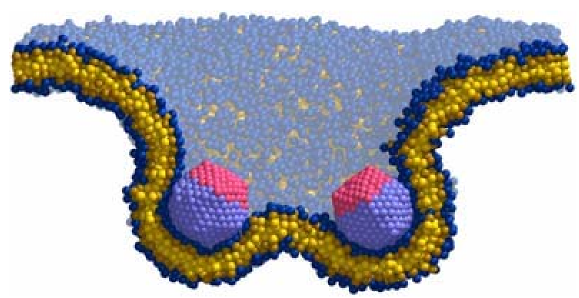
DECI Project POLYRES

Curvy membranes make proteins attractive

Cover story of Nature - May 24, 2007







a) proteins (red) adhere on a membrane (blue/yellow) and locally bend it.

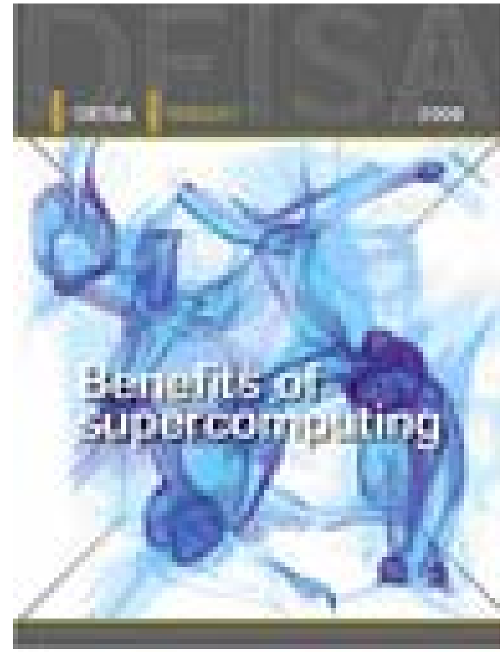
b) this triggers growing invagination.

c) cross-section through an almost complete vesicle

Benedict Reynwar, Gregoria Illya, Vagelis Harmandaris,Martin Müller, Kurt Kremer & Markus Deserno: ***Aggregation and vesiculation of membrane proteins by curvature mediated interactions***, NATURE Vol 447, 24 May 2007

Distributed European Infrastructure for Supercomputing Applications

Read more in:



Thank You!

Gentzsch@rzg.mpg.de