Radiation damage of biomolecules (RADAM) database development: current status

S Denifl¹, G Garcia², BA Huber⁴, BP Marinković⁵, N Mason⁶, J Postler¹, H Rabus¹, G Rixon⁶, AV Solov’yov⁸*, E Suraud⁹, AV Yakubovich⁸*

¹Institut für Ionenphysik und Angewandte Physik, Technikerstr. 25 / 3, A-6020 Innsbruck
²Instituto de Física Fundamental, Consejo Superior de Investigaciones Científicas,(CSIC) Serrano 113-bis, 28006 Madrid, Spain
³Centre for Medical Radiation Physics, University of Wollongong, NSW 2522, Australia
⁴Centre de Recherche sur les Ions, la Matière et la Photonique(CIMAP), Bd Henri Becquerel, 14070 Caen Cedex 05, France
⁵Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade and Advanced School for Electrical Engineering and Informatics, 11030 Belgrade, Serbia
⁶Open University, Faculty of Science, Walton Hall, Milton Keynes, MK7 6AA, UK
⁷Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany
⁸Frankfurt Institute for Advanced Studies, Ruth-Moufang Str. 1, 60438, Frankfurt am Main, Germany
⁹Laboratoire de Physique théorique, IRSAMC, CNRS, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse cedex, France

*On leave from A.F. Ioffe Physical-Technical Institute, Polytechnicheskaya 26, 194021, St. Petersburg, Russia

E-mail: nano-ibct@fias.uni-frankfurt.de

Abstract. Ion beam therapy offers the possibility of excellent dose localization for treatment of malignant tumours, minimizing radiation damage in normal tissue, while maximizing cell killing within the tumour. However, as the underlying dependent physical, chemical and biological processes are too complex to treat them on a purely analytical level, most of our current and future understanding will rely on computer simulations, based on mathematical equations, algorithms and last, but not least, on the available atomic and molecular data. The viability of the simulated output and the success of any computer simulation will be determined by these data, which are treated as the input variables in each computer simulation performed. The radiation research community lacks a complete database for the cross sections of all the different processes involved in ion beam induced damage: ionization and excitation cross sections for ions with liquid water and biological molecules, all the possible electron - medium interactions, dielectric response data, electron attachment to biomolecules etc. In this paper we discuss current progress in the creation of such a database, outline the roadmap of the project and review plans for the exploitation of such a database in future simulations.
1. **Introduction**

The damaging effect of ionizing radiation has been known for many years. It being commonly accepted that high-energy tracks formed by α, β, γ radiation and atomic ions ionize cell components along the track, thereby leading to various dissociation channels and to the formation of damaging radicals. This knowledge has triggered intensive research on radiation protection and the development of intensive biomedical uses of different radiation, generally called radiotherapy, used especially for tumour-based diseases. In this field, clinical trials and technological advances often develop more rapidly than our understanding of the underlying physico-chemical processes. Thus curative protocols are established on an empirical basis based on experience and expertise of the personnel involved, rather than on pure scientific knowledge.

Ion Beam Cancer Therapy (IBCT, or hadron therapy) represents a new and effective method for providing high-dose delivery directly into tumors, thereby maximizing killing of the cancer cells while simultaneously minimizing the radiation damage to surrounding healthy tissue. However, the full potential of Ion Beam Cancer Therapy can only be realized by a better understanding of the physical, chemical and biological mechanisms, on a range of time and space scales that lead to cell death under ion irradiation. The COST Action MP 1002 NANO-IBCT aims to combine, using a multiscale approach, the unique experimental and theoretical expertise available within Europe to acquire a greater insight into radiation damage induced by ion impact at the nanoscopic and molecular level. For more information, visit [http://fias.uni-frankfurt.de/nano-ibct](http://fias.uni-frankfurt.de/nano-ibct).

2. **Current Status**

Theoretical investigations of the effect of ionizing radiation are usually conducted using computer simulations, since the underlying molecular processes are too complex to treat them purely analytically. Several computer codes for radiation transport (e.g. Geant4: [http://geant4.cern.ch/](http://geant4.cern.ch/), FLUKA: [http://www.fluka.org/](http://www.fluka.org/), MCNP: [http://mcnp-green.lanl.gov/](http://mcnp-green.lanl.gov/)) and track structure analysis (KURBUC, PARTRAC, LEPTS, OREC, SHERBROOKE) have been developed for the simulation of interactions of ionising radiation with various media. MBN Explorer [1] ([http://www.mbnexplorer.com/](http://www.mbnexplorer.com/)) is a multi-purpose computer code, which is designed to study molecular systems of various degrees of complexity. A broad variety of interatomic potentials implemented in the MBN-Explorer allows a simulation of the structure and dynamics of different molecular systems, such as atomic clusters, fullerenes, nanotubes, proteins, DNA, composite systems, nanofractals, etc. A recent project exploits MBN-Explorer to focus on the thermo-mechanical pathways of DNA damage as a consequence of irradiation in ion beam cancer therapy [2,3].

However, such simulations are strongly dependent upon the quality of the input data. At a recent workshop organized by the COST Action Nano-IBCT in Vienna February 24-26, 2012 the need for a new database for the data on the processes relevant for the interaction of ionizing radiation with organic molecules became obvious. It was decided that this database entitled RADAM should exploit the experience of the VAMDC project [http://www.vamdc.eu/](http://www.vamdc.eu/), which is devoted to the creation of an interoperable e-Infrastructure for the exchange of atomic and molecular data. The RADAM database could be extended with the data from simulated outputs (using a database network model), to compare the capabilities of different computer codes.
The general structure of the RADAM Database is to consist of five principal areas: Ionic Interactions, Electron/Positron Interactions, Photonic Interactions, Multiscale RADAM phenomena and Radiobiological phenomena, see figure 1.

Each of the aforementioned areas will be represented by at least one database node, where the corresponding information will be stored. Data stored at all the nodes should be accessed through a single user portal (RADAM DB portal), which will provide the user interface for creating database queries using a unified query language for all the databases. This query language is currently the VAMDC SQL subset 2 (VSS2), but might be extended in the future as necessities arise.

The RADAM DB nodes will be also connected to the VAMDC portal in order to provide the possibility for users to search for data compatibility with other VAMDC databases.

Currently the following nodes are being created for the RADAM Database:

2.1. Cross sections for dissociative electron attachment

The Nano-Bio-Physics group at the University of Innsbruck has established a database node containing cross sections for dissociative electron attachment (DEA) processes in the gas phase. The node is called “Innsbruck DEA Database” (IDEADB) and is fully compatible with VAMDC standards version 11.12 (the current, production release) and 12.07 (the forthcoming release). Additionally, a simple web interface was developed in order to allow easy inspection and comparison of data (http://ideadb.uibk.ac.at)

IDEADB currently contains gas phase measurements of the electron energy dependence of cross sections for DEA processes to various molecules, some of which are biomolecules. For each process, the originating species and resulting anion fragments are described in as much detail as possible (e.g. InChI for molecules where possible) and the electron energy resolution used for the measurement is given along with comments taken from the original publication. Since all data is taken from publications and DOIs are given for each process, the data can be cited easily and the additional information given by a full publication can be accessed. For further investigations of data, each cross section can be exported to an ASCII file.

The system is based on the Python implementation of the Node Software developed by the VAMDC project. It uses nginx as a reverse proxy and gunicorn as an application server to power the Django framework. Since the amount of data is rather small, the VAMDC performance goals can be met easily. The system is hosted on dedicated hardware at the University of Innsbruck and therefore well connected to scientific networks.

2.2. Electron and positron interactions with biomolecules
2.2.1 Low energy (0-10 keV) electron and positron interaction
Low energy (0-10 keV) electron and positron interaction cross sections and energy loss in biomolecular systems are being collated in a database prepared by G García and co-workers at CSIC Madrid. Cross section data for electrons and positrons derived from the IAM-SCAR [4,5] calculation procedure are stored together with a critical compilation of available theoretical and experimental data. This node is connected with the Centre for Antimatter-Matter Studies (CAMS, Australia) and the Laboratory for Atomic Collisional Processes of Institute of Physics University of Belgrade (see subsection 2.2.2).

2.2.2 Electron/positron interactions with biomolecules
Electron/positron interactions with biomolecules node is created in Belgrade at the Laboratory for Atomic Collisional Processes of Institute of Physics University of Belgrade, IPB (mail.ipb.ac.rs/~centar3/acp.html). Primarily data measured at the own laboratory will be stored in the node. Those include absolute electron differential (in angle and energy) cross sections for elastic scattering and excitation of biomolecules and molecules of biological interest, but will also include targets such as metal atoms. Secondary data sets will be assembled by collating calculated data for the same target measured at IPB. All data are in compliance with VAMDC standards and the xsams scheme for the collisional processes. Data sets are presented as 3D surfaces (cross sections versus electron impact energy and scattering angle). Data error analysis is included as absolute error bars for each data point.

2.3. Ion interactions with biomolecules and clusters
The CIMAP group in Caen intends to create a node in which data concerning the interaction of singly and multiply charged ions with molecules and clusters of biological interest will be stored. Initially cross section data for the elastic and inelastic (ionization and charge transfer) scattering of ions from biomolecular systems are being collated. These data will be presented as a function of the scattering angle and the collision energy in tabular and in graphical form. Targets will include small biomolecules such as halogenated nucleobases, nucleosides and nucleotides, amino acids and (poly) peptides as well as clusters of these biomolecules. Where possible, experimental data will be compared with theoretical results, for example as presented for uracil in [6]. Furthermore, the RADAM database will include ion-induced relative fragmentation yields as a function of the charge state and kinetic energy of the projectile (see for example [7]). In a second step, the database will be extended to include information on secondary particle production (e.g. electrons and radicals) induced by ion impact.

2.4. Interaction of ionizing particles with all constituents of DNA
In an ongoing research program at PTB, cross section data for the interaction of ionizing particles with all constituents of the DNA or of model molecules for these building blocks (such as tetrahydrofuran for deoxyribose or pyrimidine for the nucleic bases cytosine and thymine) are determined. For electrons with kinetic energies between 20 eV and 1 keV, total scattering cross sections as well as single differential elastic and double differential inelastic cross sections are determined, where the experimental apparatus covers a range of scattering angles between 15° and 135° [8]. The experimental data are extrapolated to forward and backward scattering angles via phenomenological models that are supported by theoretical calculations based on, for example, the modified independent-atom model [9]. Total elastic and single differential inelastic electron cross sections are calculated by numerical integration, and the latter are interpolated as a function of energy using the binary-encounter Bethe model [10].
A second experiment that determines the absolute cross sections for the fragmentation of biomolecules after impact ionization has recently started operation and is currently used for measurements of
electron impact on the DNA molecular building blocks tetrahydrofuran and pyrimidine. The sample preparation system is based on a specially designed supersonic jet that allows also non volatile substances like the nucleic bases to be studied.

A third experiment for measuring double differential ionisation cross sections of DNA constituents for impact of protons and light ions with energies typical for the Bragg-peak region close to the end of the ion track is under construction. This experiment will also allow the determination of cross sections for charge-transfer between these projectiles and target molecules. The data will be evaluated in terms of analytical models for the energy dependence of the cross sections such as the HKS model [11].

As an input to the data base, the following data will be provided: the experimental values including an uncertainty estimate; a report on the evaluation for consistency; and description of the parameters of the evaluated models with an uncertainty estimate for this best-fit interpolation that takes correlations into account.

2.5. **Multiscale RADAM phenomena**

The Multiscale RADAM phenomena node has been created in Frankfurt in the Meso-Bio-Nano Science group of Prof. Dr. Andrey Solov'yov (fias.uni-frankfurt.de/mbn). The node will store experimental and theoretical data on RADAM related phenomena such as stopping power, diffusion coefficients, Bragg peak characteristics, thermo-mechanical damage, beams defocusing/spreading, complex dielectric functions, molecular fragmentation cross sections, chemical reaction rates. See subsection 4 in the appendix for details. The initial version of Multiscale RADAM phenomena node was created in collaboration with the group of Prof. Bratislav Marinković at Institute of Physics, Belgrade. Currently the node stores experimental data on the linear energy transfer for carbon ions with different initial energies in water. The corresponding links to the publications, where the data was reported, are provided for each query.

3. **Roadmap of the RADAM DB project.**

The roadmap of the RADAM DB project is presented in figure 2. In order to create the database it is necessary to define its principal logic structure. Working group leaders responsible for each database component have proposed structures for each of the topical areas and these are shown in the appendix.

![Figure 2](image)

**Figure 2.** The roadmap of RADAM DB project. Creation of the database implies elaboration of the following questions: structuring, validation of the data and creation of the software, physical location of the nodes, their maintenance.
At the current stage of the development we will include only data, which has published in peer-reviewed journals. However, in certain cases unpublished but thoroughly checked data can be also of high interest to the community. We plan to develop protocols for data verification at later stages of the project in collaboration with the SUP@VAMDC project (www.sup-vamdc.vamdc.org).

4. Conclusions and future plans
The RADAM database being created within the framework of the COST Action Nano-IBCT will have a multi-node expandable infrastructure. Each node will be devoted to a certain RADAM data domain. Several nodes devoted to similar data domains can be maintained by different groups.
A dedicated RADAM data portal for constructing queries to the nodes will be created. The RADAM portal will support all types of data stored at the nodes. The nodes will be also accessed by the queries from the VAMDC portal. In case of data compatibility with VAMDC XSAMS standards the reply to the VAMDC queries will be created. Further modifications of VAMDC SQL subset will include extensions towards broader coverage of RADAM data by VAMDC queries.
Any group willing to provide data for RADAM DB or setting up a personal node is cordially invited to participate in the project. RADAM DB developers within the COST Action will provide the joining groups with necessary software for the node installation as well as help and support on configuration and maintenance of the node. The RADAM DB developers can be accessed by e-mail to nano-ibct@fias.uni-frankfurt.de or by direct contact with the working group leaders.

5. Acknowledgements
Parts of this work were supported by the Fonds zur Förderung der wissenschaftlichen Forschung (FWF), I978 and 22665, Wien.
AY, AS and BM acknowledge the support from bilateral project between Germany (DAAD ID 54394840) and Serbia (#680-00-00095/2012-09/9; OI 171020). BM acknowledges the support from VAMDC.
The authors also acknowledge the support and inspiration from COST Action 1002 Nano-IBCT under whose auspices the RADAM DB has been developed.

References
[1] Solov'yov IA, Yakubovich AV, Nikolaev PV, Volkovets I, Solov'yov AV 2012 J. Comp. Chem. 33 2412
[2] Yakubovich AV, Surdutovich E, Solov’yov AV 2011, AIP Conf. Proc. 1344 230
[3] Yakubovich AV, Surdutovich E, Solov’yov AV 2012 Nucl. Instr. Meth. B 279 135
[4] Blanco F and Garcia G 2003 Phys. Rev. A 67 022701
[5] Blanco F and Garcia G 2004 Phys. Lett. A 330 230
[6] Champion C et al. 2012 JPCS 373 012004
[7] Maclot S. et al. 2012 JPCS 373 012004
[8] Bernal MA and Liendo JA Nucl. Instr. Meth. Phys. Res. B 2006 251 171
Appendix. RADAM DB structure and topical areas

1. Ionic interactions

   Ionic interactions

   - Cross sections for ion - BM interactions
     - Elastic scattering CS
     - Ionization CS
     - Ionization energies
     - Charge transfer CS
     - Molecular fragmentation CS
   - Molecular fragmentation (direct processes)
     - Fragmentation spectra
     - Relative fragmentation yields
     - DNA strand breaks
     - Fragmentation mechanisms
   - Secondary particle production
     - Secondary electrons (yields, energy spectra)
     - Radical production and reactions

2. Electron/positron interactions

   Electron/positron Interactions

   - Cross sections for e+/e- - BM interactions
     - Elastic scattering CS
     - Ionization CS
     - Ionization energies
     - Excitation CS
     - Positronium CS
     - CS at resonances
   - Molecular Fragmentation
     - DEA spectra
     - Relative fragmentation yields
     - DNA strand breaks
     - Fragmentation mechanism
   - Secondary particle production
     - Yields of exc. or ion., particles energy spectra
     - Radical production and reactions
     - Annihilation
     - Positronium formation
3. **Photonic interactions**

   ![Photonic interactions diagram]

4. **Multiscale RADAM phenomena**

   ![Multiscale RADAM phenomena diagram]

5. **Radiobiological phenomena**

   ![Radiobiological phenomena diagram]

To be elaborated