DMP of project "Bio-relevant anion radicals excited states dynamics study in solution by pump–repump–probe spectroscopy"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "Science Europe : modèle structuré" fourni par Science Europe.

Renseignements sur le plan

Titre du plan	DMP of project "Bio-relevant anion radicals excited states dynamics study in solution by pump–repump–probe spectroscopy"
Version	Version intermédiaire
Domaines de recherche (selon classification de l'OCDE)	Chemical sciences, Physical sciences
Langue	eng
Date de création	2023-01-03
Date de dernière modification	2023-01-03
Identifiant	RAD3P_DENISOV_ICP
Type d'identifiant	Identifiant local
Documents (publications, rapports, brevets, plan expérimental….), sites web associés	• web-site : https://elyse-platform.academy/

Renseignements sur le projet

Titre du projet	Bio-relevant anion radicals excited states dynamics study in solution by pump–repump–probe spectroscopy
Acronyme	Rad3P

Résumé	The objective of the Rad3P project aims to study for the first time the evolution of excited states of transient anion radicals of DNA/peptides models (quantum efficiencies of state-to-state transition) in aqueous solution. Excited-state chemistry of anion radicals is important in the light of dissociative electron attachment (DEA) since it could play role in reductive damage path for DNA/peptides through the formation of the excited anion radicals undergoing irreversible dissociation. In the course of the project Rad3P, unique radiolysis spectroscopy will be developed based on “Pump-Repump-Probe” spectroscopy approach, by replacing Pump with an electron Pulse, what was never performed in radiation studies. The “Pump(electron pulse)-Repump(laser)-Probe(supercontinuum)” spectroscopy consists of reexciting solute anion radicals formed in reaction with descendants of secondary electrons: presolvated and solvated by a short (100 fs) optical pump (variable energy). The fate of excited anion radicals could be then studied with a proper time resolution with a third spectrally broad 100 fs probe pulse, namely following bleaching of ground-state absorption of the radical anion and the dynamics of excited-state evolution. The newly generated knowledge will have great importance to radiotherapy techniques fighting cancer, to the codes devoted to the modeling of ionization radiation interaction with matter, DNA/peptide damage which ignores DEA at the moment. In addition, the project addresses both technological and scientific challenges by developing new spectroscopy extending the applicability of pulse radiolysis technique. It will provide a new way for research under ionization radiation since it is clear that there is plenty of room in the short timescale. During this project, we will also implement a data management plan to give access to the community the generated experimental data.
Sources de financement	• French National Research Agency : ANR-20-CE29-0003
Date de début	2021-01-01
Date de fin	2024-12-31
Partenaires	• Institut de Chimie Physique (200011673T)

Produits de recherche :

1. Transient absorption data (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
DENISOV Sergey	Insitutute de Chimie Physique	• Coordinateur du projet • Personne contact pour les données • Responsable de la conservation à long terme des données • Responsable de la documentation des données • Responsable de la production ou de la collecte des données • Responsable de la protection des données • Responsable des questions éthiques • Responsable du dépôt et de la diffusion des données • Responsable du plan de gestion de données • Responsable du stockage des données • Responsable du traitement et de l'analyse des données • Responsable juridique

DMP of project "Bio-relevant anion radicals excited states dynamics study in solution by pump–repump–probe spectroscopy"

1. Description des données et collecte ou réutilisation de données existantes

1.1 Description générale du produit de recherche

Nom	Transient absorption data
Description	The data represents the transient absorption data of studied systems using pump-repump-probe spectroscopy.
Type	Jeu de données
Mots clés (texte libre)	
Langue	eng
Identifiant pérenne	file_name_project_name_id
Type d'identifiant	Identifiant local
Contient des données personnelles ?	Non
Contient des données sensibles ?	Non
Prend en compte des aspects éthiques ?	Non

1.2 Est-ce que des données existantes seront réutilisées ?

Justification	For a particular project, it is unlikely that any data will be reused since the new data will be generated and analyzed. However, for calibration purposes, the existing data in the literature or in local storage ould be resued

1.3 Comment seront produites/collectées les nouvelles données ?

Titre de la méthode	time-resolved spectroscopy
Description	The data will be collected using LabVIEW and Python written software, which will gather the intensity of light passing through the sample, recorded as spectra at different times of delays between excitation pulses, working parameters of accelerator (variation of dose per pulse; currents; vacuum variation) and laser (variation of energy per pulse). The absorption spectrum of the solvated electrons in water and other solvents will be utilized as a standard for the correction of obtained transient spectra. All recorded experimental results will be stored in HDF format, which will include information about the project, its goals, expected results, references to the literature, and analyzed data, i.e., metadata.
Equipements, plateaux techniques utilisés	• ELYSE platform : linear electron accelerator dedicated to pulse radiolysis

2. Documentation et qualité des données

2.1 Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

Description	The metadata format that is planned to be used will follow the descriptive format of Dublin Core. Data Storage: The HDF5 file format allows easily storing strings of variable length, thus it allows easily to store within the file all required information (metadata) about the particular project's measurement: HDF5 (Project, ID, goal; important references and links, creator ID): -measurement group1..N (ID, goal; important references and links, creator ID): - different measurement dataset (operator ID; ID of measurement; observations; who created; who did changes; -treatment group1..N (ID, goal; important references and links, creator ID). The RAD3P project file structure that will be stored on a specially dedicated hard drive will include following folders: - Folder #Description (includes files describing the project, main schemes, and major orders. hardware manuals and drivers, software, etc.) - Folder #Data - Folder #Papers - Folder Students (Includes Ph.D. and Master student information) - Folder #Related Literature and Links
Code langue des métadonnées	eng

2.2 Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

3. Exigences légales et éthiques, code de conduite

3.1 Quelles seront les mesures appliquées pour assurer la protection des données à caractère personnel ?

Description	No personal data will be processed except the name of operators and project holders.

3.2 Quelles sont les contraintes juridiques (sensibilité des données autres qu'à caractère personnel, confidentialité, ...) à prendre en compte pour le partage et le stockage des données ?

Description	The data belong to the CNRS, which defines intellectual property (IP) politics in agreement with the other institutional partner University Paris-Saclay. Whenever possible and meaningful, results will be submitted to institutional IP offices for possible patent filling. Otherwise, data will be shared under the Creative Commons Attribution Non-Commercial Share Alike license (CC-BY-NC-SA).

3.3 Quels sont les aspects éthiques à prendre en compte lors de la collecte des données ?

Description	We follow CNRS' guidelines for conducting research responsibly, following the ethics code.

4. Traitement et analyse des données

4.1 Comment et avec quels moyens seront traitées les données ?

Description	Home-written code that will be available trought the GitHub repository.

5. Stockage et sauvegarde des données pendant le processus de recherche

5.1 Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Besoins de stockage	The data that is collected on the ELYSE platform every day is copied automatically and monthly the data is transferred on the external drive. The whole dataset will be stored in HDF format files, e.g. HDF5, that will contain the multidimensional arrays of raw spectra (numeric data) of supercontinuum (with electron pulse, without electron pulse, with and without optical repump) both for signal and reference channels. The expected volume for a single measurement is estimated to be 10 Mb, taking into account, that it requires a minimum of 10 scans for good statistics, it would take 100 Mb for a sample, including parameters of the accelerator and laser not exceeding 1 Mb. The rough estimation for a project would be ca. 5Gb. Considering that the RAD3P project will include several different molecular systems to study, i.e, separate projects, the total volume for data storage that will be required does not exceed 200Gb for raw data. However, to be realistic, 10 times more is needed for storing analyzed data, which gives a lower boundary of 2000 Gb for 36 months of the RAD3P project. In order to facilitate the re-utilization of the data, the python-based software will be shared.
Volume estimé des données	2000
Unité	Go
Mesures prises pour la sécurité des données	The collected data will not have any sensitive information. However, all hard drives in the laboratory are crypted. The excess to raw data will be restricted to members of the team before publishing the results. Later the data will be available for download on-demand. The list of available data will be presented on the elyse-platform.academy/projects website. In order to recover data from events such as fire, the external drive will be held separately in another place in the same laboratory. The most crucial data related to the main results of the project will be stored in the CNRS storage cloud. Data security is handled by the institutions CNRS and University Paris-Saclay.

6. Partage des données et conservation à long terme

6.1 Comment les données seront-elles partagées ?

Modalités de partage	Only data selected for scientific publications will be made available in the long term using a trustworthy repository (Zenodo) as pre-treated raw and raw data. Other data will be available on demand. For publication mainly open-access journals will be selected. However, preprints of the publications will be made available in HAL (according to the French legislation). The access to raw experimental data will be regulated by IP and will be provided in the open format followed by metadata to enable re-use.
Potentiel de réutilisation	for data verification
Entrepôt/Catalogue de données	• HAL : https://cat.opidor.fr/index.php/HAL

6.2 Comment les données seront-elles conservées à long terme ?

Justification	The holder of the project and the hired Ph.D. will devote a part of their time to publishing data on public repositories while writing the publications. The long-term storage will be ensured by the free service provided by Zenodo (up to 50 Go), and double storage on local hard drives. The whole dataset will be stored in HDF format files, e.g. HDF5, that will contain the multidimensional arrays of raw spectra (numeric data) of supercontinuum (with electron pulse, without electron pulse, with and without optical repump) both for signal and reference channels. The expected volume for a single measurement is estimated to be 10 Mb, taking into account, that it requires a minimum of 10 scans for good statistics, it would take 100 Mb for a sample, including parameters of the accelerator and laser not exceeding 1 Mb. The rough estimation for a project would be ca. 5Gb. Considering that the RAD3P project will include several different molecular systems to study, i.e, separate projects, the total volume for data storage that will be required does not exceed 200Gb for raw data. However, to be realistic, 10 times more is needed for storing analyzed data, which gives a lower boundary of 2000 Gb for 36 months of the RAD3P project. In order to facilitate the re-utilization of the data, the python-based software will be shared.
Volume estimé des données	50
Unité	Go
Date de début	2021-01-01
Date de fin	2024-12-31