JNCS
Journal of Nepal Chemical Society
March/April 2023, Vol. 43, No. 2, 11-22
ISSN: 2091-0304 (print)
DOI: https://doi.org/10.3126/jncs.v43i2.53337
RESEARCH ARTICLE

Free and Open Source Codes for Computational Chemistry Research Initiation: A Conspectus

Jhashanath Adhikari 'Subin'[1, 2, *], Rameshwar Adhikari[2]

[1]*Research Center for Applied Science and Technology (RECAST), Tribhuvan University, Kirtipur, Kathmandu 44600*

[2] *Computational Material Science Division, Scientific Research and Training Nepal P. Ltd., Bhaktapur 44800*

** Corresponding email: subinadhikari2018@gmail.com*

Submitted: 17 July 2022, Revised: 31 Aug 2022, Accepted: 7 sept 2022

Abstract

Computational chemistry is a field of current active research in science that can supplement and complement the regular experimental methods and validate or propose new theoretical formulations. It has contributed from basic to advanced levels of investigation in a synergistic manner. The mathematical experiments can be pursued with minimal resources and is sustainable making it a good approach for the institutions with limited resources. The freely available computer programs that are suitable for solving various types of problems in chemistry and interdisciplinary areas are briefly mentioned. The general features and specificity of some selected codes that are adaptable for undergraduate and graduate level courses and research in Nepal are highlighted. Recommendation is made for the incorporation of adequate theoretical courses and parallel laboratory sessions in the existing syllabus for obtaining trained and skilled work force qualified for national and global careers in computational chemistry and also for research and development in Nepal.

Keywords: computational chemistry, *ab initio*, free and open source softwares, *in silico,* first principles

Introduction

Computational chemistry uses computing power to solve theoretical equations representing the physical world, to explain the chemical phenomenon and to determine the chemical properties. The computational methods have become indispensable tool in modern scientific research that in itself is a standalone field and also in collaboration with various types of experimental methods thus spanning from pure theoretical work to advanced interdisciplinary areas [01]. Apart from testing the theories, not only has it solved the abstract problems in chemistry and related areas in an independent manner but also has provided unprecedented insights into undoable procedures making the combination tenable with synergistic advantages [02]. The relationship between theory, experiment and computation can be analogous to a three-wheeled cart with each wheel guiding the other two in synchronicity to reach a common solution as shown in Figure 1.

In context to our country, the fundamental investigation in diverse areas like material modeling, surface science, and pharmacology that can be carried out by computational tools seems important and feasible. The search with phrases "computational chemistry", "*ab initio*", "first principles" and "*in silico*" in google scholar returned *ca.* 18k, 24k, 33k and 41k results respectively for the time period starting Jan 01, 2021 up to July 13, 2022. It indicates significant research activity in the global community and also the popularity of the method.

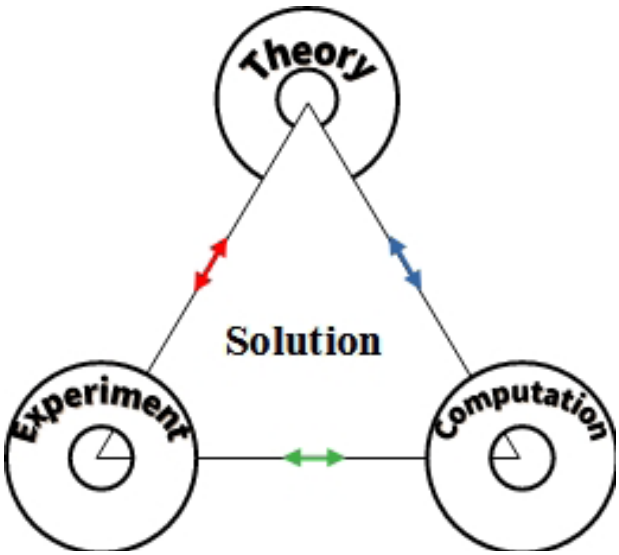


Figure 1: *Schematic diagram of a three-wheeled cart showing interrelationship between theory, experiment and computation; the double sided arrow head represents that each interacts with the other two to drive to a better solution*

There are diverse methodological tools available to address multitude of problems considering the broadness in chemistry and deepness of each specialty area [03]. The proper understanding in terms of resolution of distance (ranging from the nucleus and atomic level up to the whole macro-scale crystal) and in terms of time (from 0.5 femtosecond to second and up) can be achieved employing different methods of calculations [04-05]. The research problems based on national priorities and personal interests can be tailored to be solved by utilizing appropriate computational programs that are freely available [06-07]. The results can be correlated or validated with those obtained from experiments for better and strong inferences.

Technical knowledge is not in par with the theoretical understanding because of the lack of adequate computational courses and associated laboratory sessions in chemistry curriculum in Nepal [08]. This conspectus points out to the availability of several free and open source softwares (FOSS) and capability of each one in specific areas with the intention of providing quick and concise information to the faculty and students interested in computational chemistry or material science research [09]. Basis for recommendation will be constructed highlighting the present research trends and the need of theoretical knowledge and skills required by the current generation to participate in it using different computational tools and techniques.

Computational Methods- An Overview

Depending upon the nature of problem to be solved, parameters to be determined and the resolution to be attained, appropriate computational method is to be selected. There are multitude of programs available each with different methods as modules or sub-routines and with different approximations. Broadly speaking, four domains of system in terms of length can be studied and the corresponding time resolution of the phenomenon can be obtained by different methods as shown in Figure 2.

The regularly used terminology *ab initio* or first principles in this field, refers to calculations from the start or beginning without any usage of experimental results. But most of these computational methods use published structures (X-ray or neutron diffraction, CryoEM derived) as the starting point of the calculation. Hence this misnomer requires to be revised to second principles or experimental structure calculations. Most of the open source codes discussed here use the space group setting [10], atomic coordinates and lattice parameters of the crystal derived from experimental procedures. Hence some knowledge on solid state or crystallography seems to be helpful in creating the input files for setting up a calculation.

The basic chemistry or physics at the atomic and molecular level (picometer, femtosecond) can be solved by quantum mechanics with high accuracy and by semi-empirical methods with reduced accuracy. The smaller the distance and shorter the time scale to be measured, more resources are required for simulating the material and for approximating its properties. There is always a trade-off between computing-time and accuracy which suggests the selection of optimized resources and method for justifiable results [11].

The quantum mechanical methods include Hartree Fock, post Hartree Fock, density functional theory,

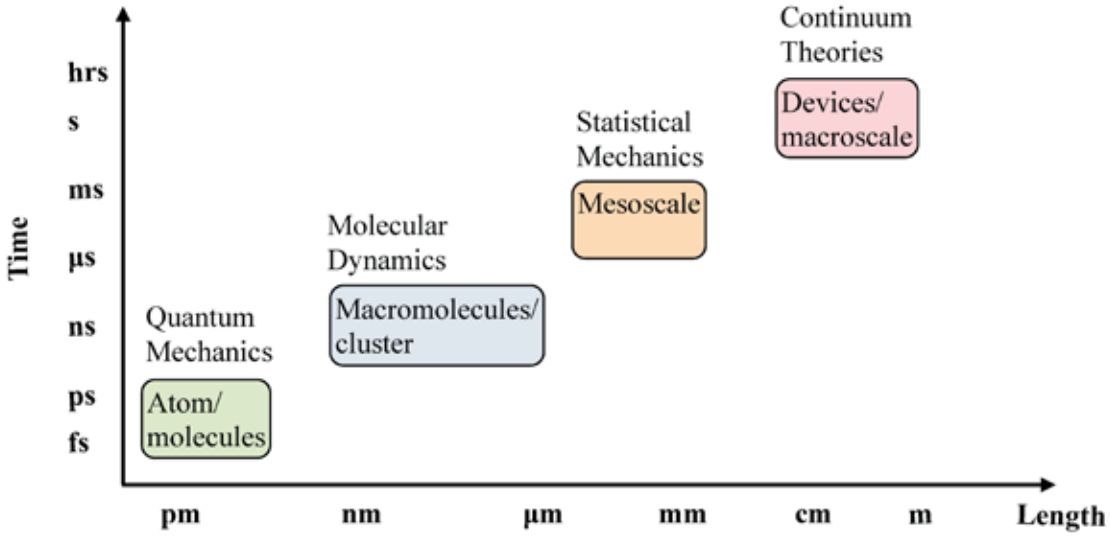


Figure 2 *Size and time scales with computational methods of different systems; Quantum mechanics and molecular dynamics are two popular choices for many problems in chemistry*

second order Moeller Plesset perturbation theory, coupled cluster singles and doubles etc. that can be applied in the ascending levels of theory [12]. Also stands the selection of representation of elements or group of atoms a tedious task in solving time (in) dependent Schrödinger wave equation or Kohn-Sham equations and beyond. The system can be represented to be composed of simple atoms or cluster of atoms limiting to isolated truncated molecule or to be the whole crystalline solid using repetition of unit cell with periodic boundary conditions. More elaborately, atomic orbitals, localized basis sets (Pople, Dunning, Jorge, Karlsruhe, Jensen etc), pseudo potentials (NCPP, USPP), projected augmented waves, linear augmented plane waves etc can be chosen depending upon the type of chemical system and the desired property to be determined [13-14]. The exchange correlation functionals (LDA, GGA, meta and hybrid), dispersion interactions, solvation and spin orbit coupling also constitute major components of the computational parameters [15]. The material (metallic, semi-conductor or insulator) in zero dimension (atom, nanoparticle), one dimension (nanorod and nanotube), two dimensions (surface or nanosheet) and three dimensions (bulk solid) can be studied by suitable programs.

One of the recognizable contribution of quantum mechanical calculations that is worth mentioning is the development of non-platinum based oxygen reducing catalysts. An alternate material made up of monolayer boron nitride (BNNS) on metal substrate as shown in Figure 3 was proposed to possess comparable activity based on the density of states whereby the energy gap was reduced from 4.6 eV to semiconducting region due to the extension of electronic states towards the Fermi level [16]. It was later proven experimentally and glassy carbon substrate in place of Au(111) was taken as the control. This highlights the guiding nature of computational method in designing new type of materials.

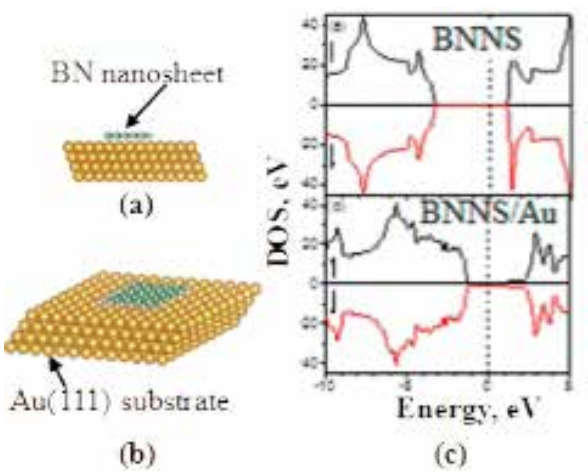


Figure 3: *Schematics of Boron Nitride Nanosheet (BNNS) over Au(111) surface (a) side view (b) birds-eye view and (c) the DOS plots. Boron atom as green sphere, nitrogen as blue sphere and gold as yellow sphere.* ***(Reprinted with permission from*** *J. Am. Chem. Soc. 2014,* ***136****, 18, 6542–6545 Copyright 2014 American Chemical Society)*

For bigger systems like macromolecules and clusters (nanometer, nanosecond), classical methods using

molecular mechanics (Lennard-Jones potential) is the easiest and the fastest way to characterize. The consideration of point charges for atoms and springs for bonds makes feasible the study of a large collection of atoms. Others approaches like Monte Carlo, molecular dynamics, semi-empirical approximations are also used for large molecular systems. Nanoparticles and biomolecules are basically investigated by this approach. The proteins and peptides are studied and especially the binding of small ligands to the active sites for its inhibition (structure based drug design) is estimated. The stochastic search method (molecular docking) that leads to the best docked complex [17] and assessment of its stability by molecular dynamics simulation (different force fields) stands as another category of computational research [18]. The homology modeling (used in ligand based drug design), protein-protein, protein-DNA etc interaction study comprises active research in bioinformatics and computer aided drug design (CADD).

Mesoscale and macroscale (engineering systems) systems along with online server based research, machine learning, artificial intelligence (AI), deep learning and post-processing or visualization tools will not be dealt in this conspectus due to its limited scope as a primer.

Since our country is rich in natural resources (water, minerals and herbs), utilization of these substances to produce value added products would be beneficial. The research problems concerning these materials that requires computational tools and procedures to provide solutions and to create a knowledge-based recommendation for further experimental work would be highly beneficial. This will pave a way for transferring the technology to trade and in the utilization of domestic resources. Since skilled researchers are required for this category of work, incorporation of computational and associated courses in the current syllabus with the motive of training the student body with the appropriate theoretical understanding and technical skills seems relevant.

The readily available FOSS that can calculate different properties of various types of systems will be dealt in next.

FOSS and their features

There are a lot of computer programs available with different types of licensing (free, academic, commercial and governmental). The open source codes (GNU General Public License) targeting specially the academia has been on the rise because of their robustness and the active global community ready to support it in terms of development, maintenance and upgradation. Each program has its own pros and cons with the following selected ones (not to be limited to) being discussed in terms of their salient features and are deemed suitable for research in our country at both undergraduate and graduate levels.

Abinit [19]. It has broad range of functionalities and is aimed for beginners with explicit tutorials to experienced users with advanced topics. It can be used to calculate the structural, electronic, magnetic, elastic, transport, excited state and optical properties along with nuclear magnetic resonance (NMR) parameter, the electric field gradient (EFG) tensor. It also uses density functional perturbation theory for calculating response functions. It can perform molecular dynamics (MD) simulation and can calculate Wannier functions. The system that can be studied are simple molecules, surfaces and solids with easy handling of input files. The post-processing tools make the data analysis simple and the thorough tutorials make the whole program quickly learnable. The code can run parallel with GPU acceleration and input file generators are available. It is incapable of performing full-potential all-electron augmented plane wave (APW) calculations. Some properties like chemical shielding anisotropy tensor (CSA) and g-factor of electron paramagnetic resonance (EPR) are, however, missing. Preliminary knowledge on python programming is beneficial for post-possessing of the output data.

Website: https://www.abinit.org/

Quantum Espresso (QE) [20]. It has similar features to that of Abinit and is also aimed from beginners to advanced users. It can be used to calculate the structural, magnetic, elastic, electronic, transport, excited state and optical properties along with NMR parameters, the EFG tensor and chemical shielding

tensor. It can calculate response and spectroscopic properties. It is capable of performing CPMD and AIMD along with Wannier functions calculation. Other specific features include the calculation of g-factor of EPR, quantum transport properties and free-energy surfaces. Well documented tutorials and audio-visuals for better learning makes this program a good choice for the beginners. It is incapable of performing all-electron calculations with different wave-functions for higher level of accuracy. Input file generators are available for this program and a bash script is required for chaining different input datasets as program specific keywords are not available.

QE can perform electro-chemistry related calculations using effective screening medium and reference interaction site model methods and this feature is missing in ABINIT. In QE, the usage of more than 8 CPU-cores requires additional steps during compilation and set up of the program whereas ABINIT easily supports higher CPU count. System with up to million atoms is supported by ABINIT with the accuracy of first principle calculation (multibinit, second principles) and this feature is absent in QE. A gross comparison of QE with ABINIT shows that both the programs are similar in terms of basic properties calculation, accuracy, simulation time and control factors making them equally applicable for laboratory sessions in the curriculum or materials research.

Website: https://www.quantum-espresso.org/

CP2K [21]. It is again similar to the above two programs with incorporation of biomolecules and all-electron calculations. It is equipped with separable and non-linear core corrected pseudo potentials with some flavors of molecular dynamics. QM/MM approach is feasible along with Monte Carlo simulations. It is more suited to perform calculations on molecular systems using localized basis sets instead of plane waves designed for crystalline systems. Extensive documentation with various types of tutorials makes this program a good choice for beginners in the field. Some of the draw backs are its inability to use the GPU accelerator, to calculate the NMR parameters and to simulate the EPR spectrum.

Website: https://www.cp2k.org/

NWChem [22]. It is not much different from the above programs. It can use complete active space self-consistent field theory along with CCSD and related methods. All the NMR and EPR parameters can be determined from a single input file. It can calculate free energies of large biomolecules (QM/MM) with different force fields. Also, electron transfer calculation, vibrational SCF and quantum computing can be pursued. The GPU acceleration is limited to some parts of the program.

Website: https://www.nwchem-sw.org/

GAMESS (US) [23]. Not much different, this software requires academic licensing in addition. The tutorials are well documented and references are provided for all the key ingredients of the program. Apart from DFT, it can perform calculations employing configuration interaction, second order perturbation theory and coupled-cluster methods. Large systems can be handled as the computation is divided into smaller fragments using fragment molecular orbital approach. Also, excited states are dealt-in accurately with equation of motion and time-dependent DFT methods. It is a good alternative to CP2K, NWChem and other similar programs.

Website: https://www.msg.chem.iastate.edu/gamess/

LAMMPS [24]. It is a program that can model an atom and macroscopic systems with billions of atoms by using classical molecular dynamics. It can perform Monte Carlo and direct simulation Monte Carlo for fluids. Different MD simulations like path-integral, targeted, steered etc can be performed. It can be GPU accelerated and contains numerous additional capabilities. The code is well documented with thorough and step-by-step tutorials. It is not an electronic structure code and is suited for beginners to advanced users for studying inorganic and biological systems in terms of time dependent properties.

Website: https://www.lammps.org/

SIESTA [25]. It is an electronic structure and molecular dynamics program for molecules and solids. It uses finite-support pseudo atomic orbitals and NCPP with real-space grid for wavefunction and

density components. Quantum electronic transport and dielectric polarization can be studied with it. The program is equipped with lesser features as compared to other quantum chemistry codes and contains tutorials and visuals for beginners. It is not GPU accelerated and USPP are missing along with the module for NMR parameters determination.

Website: https://departments.icmab.es/leem/siesta/

FLEUR [26]. It is a code for performing electronic, structural and magnetic characterization of bulk, surface and linear morphologies of chemical systems using periodicity. It is based on all-electron full-potential linearized augmented plane wave plus local orbital method that is applicable to all the elements of the periodic table. The results are highly precise from the DFT calculations using this program. It can deal with Wannier functions, GW approximations, hyperfine properties and EFG tensor but cannot calculate chemical shielding. Various types of tutorials and videos lectures are present for beginners as well as for advanced learners. GPU acceleration is not present.

Website: https://www.flapw.de/MaX-6.0/

The Venn diagram (Figure 4) represents three major areas of research in chemistry with the relevant computer programs. This representation is not explicit and shows minimalistic view targeted for beginners in the field.

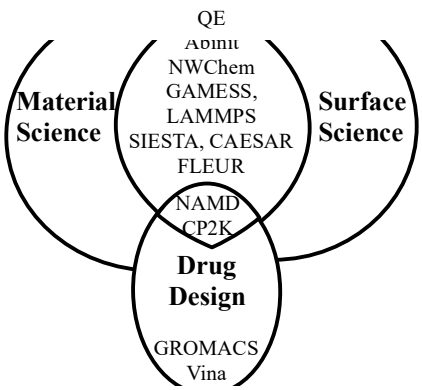


Figure 4: *Cartoon Venn diagram showing different research fields in chemistry with associated programs*

CAESAR [27]. It is a simple program for very basic users without any advanced knowledge on solid state, electronic structure theory and computer software. It gives results that are qualitative in terms of accuracy to be tested by further research. It employs extended Huckel molecular orbital method which is a semi-empirical quantum chemistry method. The crystal structures, electronic band structure, Fermi surface, electron density distribution and STM/AFM images can be calculated. Unlike other electronic structure programs, this is meant to produce quick and approximate results. The tutorial guides the users with all the basics on the theory and the working procedure.

Website: https://www.primec.com/

GROMACS [28]. It is a package for high-performance molecular dynamics simulation (*in silico*) of biochemical systems with up to millions of particles. It is GPU accelerated and contains many built-in post-processing tools. It can use multiple force fields and GUI is not present. There are many thoroughly guiding tutorials and visuals available. The protein-ligand interactions for drug discovery are best studied by this code among other different categories of research.

Website: https://www.gromacs.org/

NAMD [29]. It is similar to GROMACS code and is designed for highly scalable molecular dynamics simulations of large macromolecular systems (millions of atoms). It can perform fast calculations and is GPU enabled. It has many force fields options except GROMOS and can be automated with interactive viewing. Extensive tutorials, presentations and multimedia lectures on different problems are available and is suited for beginners as well as for advanced learners.

The data analysis requires Tcl scripting whereas in GROMACS, many parameters can be obtained directly from the built-in commands and tools. Spatial-data decomposition and load balancing makes NAMD better in handling very large cases. GROMACS tend to use efficiently only one GPU with the combination of multiple CPU cores. Nonetheless,

both the programs can be considered well suited in studying simple to moderate systems with good accuracy and reasonable time.

Website: https://www.ks.uiuc.edu/Research/namd/

AutoDock Vina [30]. It is a part of AutoDock software suite that is capable of performing stochastic search for docking of small molecules on to the macromolecular targets. The ligands are sorted out based on binding affinities derived from scoring function composed of six different parameters representing the non-covalent interactions. The procedure is documented for easy understanding and is also intended for non-computational chemists. Due to accuracy and speed, its use has been rising in the screening of thousands to millions of compounds relative to other commercial products.

Website: https://vina.scripps.edu/

Some of the programs that are popularly used in different fields of research are given in Table 1 along with the determinable features of the system studied. The general information on the fundamental as well as on the advanced level work can be obtained from the respective references provided.

Comparison of results from similar programs

In order to evaluate the performance of a program in terms of accuracy, the calculated results are compared with those obtained from the experiments. This tests the suitability of the method and of the choice of models used to describe the physical systems. Here, multiple programs with similar options and methods depending upon their availability and usability were employed to determine the NMR parameters of a three dimensional solid, $KNbO_3$. This compound contains niobium atom with quadrupolar nucleus (spin $I = 9/2$) and the crystal structure is reported in orthorhombic space group (Bmm2) settings [43]. The quadrupolar coupling constant (C_Q) and asymmetry parameter (η_Q) of ^{93}Nb nucleus are experimental observables (NMR parameters) that can also be calculated from theoretical methods. The results from free programs (Abinit and QE) and paid ones (WIEN2k [44] and NMR-CASTEP [45]) along with those from solid state NMR (ssNMR) experiments are shown in Table 2. The relaxed structure was taken into consideration for the NMR parameters simulation instead of X-ray derived structure.

The C_Q of ^{93}Nb calculated from the free programs are within 5% from those obtained from the paid softwares. The computational results are within 15% of the experimentally determined values. Since η_Q involves two principle components of electric field gradient tensor, reproducing it accurately with DFT methods is non-trivial. Here the error of at least 20% was observed implicating the requirement of stringent convergence criteria. This comparative analysis points to the equivalency of FOSS in calculating even the most sensitive parameters of a solid material. The calculation using NWChem program yielded unrealistic results owing to the use of molecular

Table 1: *Selected softwares related to research in different areas of chemistry and the studied parameters*

SN	Research areas/ special topics	Program	Observation	Refs
1.	Materials science	Abinit	Nanostructure characterization, electronic properties	[31-32]
2.	Surface science and Catalysis	QE, CP2K	Interfaces characterization, Electro-chemical implications	[33-34]
3.	Drug design	Vina, GROMACS	Phytochemical identification for cancer cells inhibition, M^{pro} inhibitor from plant based compounds	[35-36]
4.	Molecular Imaging	Siesta	STM/STS images and Raman spectra	[37]
5.	NMR spectroscopy	QE	NMR parameters	[38-39]
6.	Energy storage	LAMMPS	Nanotubes and adsorption	[40]
7.	Alloys and clusters	NWChem	Composition-structure-property relationship	[41-42]

system instead of crystalline with periodic boundary conditions. This showed that some programs may lack adequate capability of desired property determination to an acceptable accuracy. Finally, a generalization can be made in that the free programs with similar levels of calculations as in the paid versions can be easily employed for basic to advanced level of research in chemistry and material science.

Table 2: *^{93}Nb NMR parameters of $KNbO_3$ obtained from different sources*

Program/ Method	C_Q, MHz	η_Q	References
WIEN2k	-21.6	0.95	[46]
NMR-CASTEP	-19.8	0.97	[46]
Abinit	19.9	0.97	This work
QE	20.8	0.81	This work
ssNMR	22.99 ±0.03	0.822 ±0.003	[46]

Advantages of FOSS

The open source platform have made available sufficient resources to support the application of computer programs in terms of documentation, multiple tutorials, audio-visuals, bug fixes, input file generators, post-processing tools, workshops, lecture series, user support and discussion forums. The compiled codes in the form of binaries have provided quick and simple access for non-computational chemists with little or no prior experience.

Most of the free computational chemistry and physics softwares have been developed in the course of many decades with each newer version being robust than its predecessor in terms of scalability, number of features and applicability. There are many groups contributing to the addition of sub-programs or modules and fixing of the bugs in the code. The genesis of web-based libraries (git, GitLib and GitHub) for code management and sharing has made the program development transparent and easier [47]. Since the source code of the software is available freely, it can be readily modified and augmented by the global community. This has resulted in the accelerated development of FOSS that can be used in laptops to supercomputers in different operating environments. In a long run, free and open source softwares seem to be equivalent considering the prospective in performing basic to advanced level research relative to expensive commercial softwares. As an example, a commercial program VASP [48] can be considered which is similar to ABINIT and QE in terms of features, methods used and other control options. Recently, lattice parameters of different transition metal compounds have been determined using QE (GIPAW format) and have been found to be similar to those obtained using VASP (PAW format) [49]. Therefore, for the groups with limited tools, comparatively free softwares (ABINIT, QE and NWChem) or their combination seems to be appropriate, as paid version may not necessarily produce distinctly significant outcomes. Basically, same method used in different programs should produce same outcomes with identical set of input parameters and this would depict non-biased nature of the codes. Hence, in the context of the institutions in Nepal, use of FOSS would be a proper step in the initiation of computational material science research and training. Some distinct advantages of FOSS can be summarized as

a. Affordable even in a low resource settings and sustainable
b. Access to binaries ready for starters without the need of compiling
c. Detailed tutorials with step-by-step guidance in the form of manuals, slides and audio-videos
d. Capable of basic to advanced level research in different areas
e. Usable in training of computational tools and methodologies to starters
f. Freedom in adding new modules and bug fixing by global community
g. Equivalent relative to paid versions in terms of performance, capabilities and accuracy in most of the cases
h. User support as discussion forums and specialized groups (no customer support)

There are many examples of the use of the above listed programs in the literature that a quick search would reveal. Computational chemistry can thus be adopted to every instant of the occurrence of a

problem in the form of mathematical equation to be solved by it. The results would work synergistically with the experimental data for stronger justification and proper validation of an assertion.

Perspectives of future work

Considering the aforementioned research areas in chemistry and interdisciplinary fields, there are endless possibilities for computational chemistry research. A typical example that can be effectively carried out is with halloysite nanotube, an aluminosilicate clay mineral [50] as shown in Figure 5.

The material modeling involving the study of surface modification and feasibility of drug delivery can be performed using molecular dynamics (LAMMPS) and quantum mechanical calculations (QE) with this system. It can be investigated as a substrate for the possibility of adsorption of amino acids, pesticides or drug like molecules [51-52] in terms of energetics, stability and electronic structure. Hence, joining the current trend of developing new materials with novel properties using computational methods in the context of Nepal seems relevant and even more urgent.

Conclusions

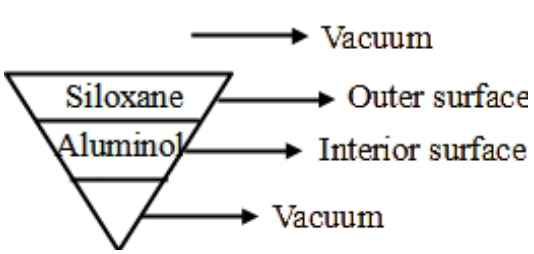
Current trend in chemistry research indicates increasing usage of computational methods for better rationalization of experimental outcomes, determination of different properties and verification of theoretical principles. The adoption of multiple approaches in solving a problem is synergistic in nature. These are supplementary and complementary to each other that lead to stronger inferences. Hence, the requirement of the inclusion of computational courses, relevant laboratory sessions and research at different levels of university syllabus has been underscored that could be tailored towards the areas of national priorities. Twelve different freely available computer programs suitable for scientific research on different aspects of material chemistry, surface science and computer aided drug design has been discussed briefly. The concepts and prospectus presented in this article is anticipated to serve as a quick guide for general understanding of the research areas covered by FOSS for solving specific types of problems therein.

Acknowledgements and Disclaimer

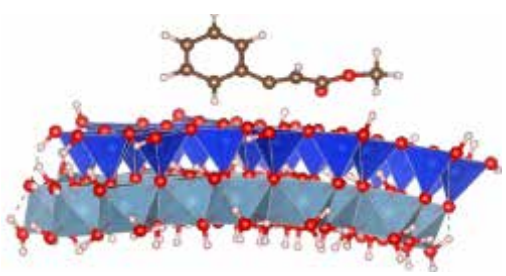
The authors would like to thank University Grants Commission (UGC) for funding (2078/2079) under the project titled "Innovative Technology Development and Materials Characterization" in the frame of "Innovative Technology Development and Scientific Exploration and Research on Nation's Priority Areas".

The helpful insights and constructive suggestions from anonymous reviewers for betterment of this conspectus is greatly appreciated.

The list of computer programs in this conspectus is not comprehensive and similar other open sources or academic codes are also available. These programs can be learned independently on a laptop computer whereas for larger systems high-performance computing is required.



(a)



(b)

Figure 5: *(a) Schematics of wedge section of spiral halloysite nanotube with outer siloxane surface and interior aluminol surface (b) Ball-and-stick model of methyl-(E)-cinnamate on top of the exterior surface of nanotube (polyhderal model, small arched section) showing the adsorption pose [50]. SiO_4 as blue tetrahedron, AlO_6 as light blue gray octahedron, carbon in dark brown, oxygen in red and hydrogen in cream.*

References

01. A. Krylov, T. L. Windus, T. Barnes, E. Marin-Rimoldi, J. A. Nash, B. Pritchard, D. G. A. Smith, D. Altarawy, P. Saxe, C. Clementi, T. D. Crawford, R. J. Harrison, S. Jha, V. S. Pande, and T. Head-Gordon Perspective: Computational chemistry software and its advancement as illustrated through three grand challenge cases for molecular science, *J. Chem. Phys.,* 2018, **149**, 180901. (DOI: 10.1063/1.5052551)

02. S. Yip, Synergistic science, *Nature Mater,* 2003, **2**, 3-5. (DOI: 10.1038/nmat778)

03. D. Young, Computational chemistry: a practical guide for applying techniques to real world problems, John Wiley & Sons. 2004, ISBNs: 0-471-33368-9

04. A. Kovalenko and S. Gusarov, Multiscale methods framework: self-consistent coupling of molecular theory of solvation with quantum chemistry, molecular simulations, and dissipative particle dynamics, *Phys. Chem. Chem. Phys.*, 2018, **20**, 2947-2969. (DOI: 10.1039/C7CP05585D)

05. D. R. Salahub, Multiscale molecular modelling: from electronic structure to dynamics of nanosystems and beyond, *Phys. Chem. Chem. Phys.*, 2022, **24**, 9051-9081. (DOI: 10.1039/D1CP05928A)

06. S. Lehtola and A. J. Karttunen, Free and open source software for computational chemistry education *WIREs Comput Mol Sci.,* 2022, e1610. (DOI: 10.1002/wcms.1610)

07. Wikipedia contributors. (2022, July 8). List of quantum chemistry and solid-state physics software. In Wikipedia, The Free Encyclopedia. Retrieved 06:07, July 13, 2022, from https://tinyurl.com/4yjyzmry

08. Bachelor's degree syllabus: https://cdctu.edu.np/bsc-syllabus-new-4-years/; Master's degree syllabus https://cdctu.edu.np/msc-syllabus-current-semester-system/ Accessed July 13, 2022.

09. In A. Metzger, Free and open source software (FOSS) and other alternative license models: A comparative analysis, 2016.

10. M. I. Aroyo, J. M. Perez-Mato, D. Orobengoa, E. Tasci, G. de la Flor, and A. Kirov, Crystallography online: Bilbao Crystallographic Server, *Bulg. Chem. Commun.* 2011, **43**(2), 183-197.

11. H. Jacobsen and L. Cavallo, Directions for Use of Density Functional Theory: A Short Instruction Manual for Chemists, In: Leszczynski J. (eds) Handbook of Computational Chemistry. 2012, Springer, Dordrecht (DOI: 10.1007/978-94-007-0711-5_4)

12. C. J. Cramer, Essentials of computational chemistry: theories and models, 2013, John Wiley & Sons.

13. N. A. W. Holzwarth, G. E. Matthews, R. B. Dunning, A. R. Tackett, and Y. Zeng, Comparison of the PAW, pseudopotential, and LAPW formalisms for density functional calculations of solids, *Phys. Rev. B,* 1997, **55**, 2005-2017. (DOI: 10.1103/PhysRevB.55.2005)

14. J. Witte, J. B. Neaton, and M. Head-Gordon, Push it to the limit: Characterizing the convergence of common sequences of basis sets for intermolecular interactions as described by density functional theory, *J. Chem. Phys.*, 2016, **144**, 194306. (DOI: 10.1063/1.4949536)

15. R. Peverati and D. G. Truhlar, Exchange–Correlation Functional with Good Accuracy for Both Structural a n d Energetic Properties while Depending Only on the Density and Its Gradient, *Journal of Chemical Theory and Computation*, 2012, **8** (7), 2310-2319. (DOI: 10.1021/ct3002656)

16. K. Uosaki, G. Elumalai, H. Noguchi, T. Masuda, A. Lyalin, A. Nakayama, and T. Taketsugu, Boron Nitride Nanosheet on Gold as an Electrocatalyst for Oxygen Reduction Reaction: Theoretical Suggestion and Experimental Proof, *Journal of the American Chemical Society,* 2014, **136** (18), 6542-6545. (DOI: 10.1021/ja500393g)

17. D. Kitchen, H. Decornez, and J. Furr et al., Docking and scoring in virtual screening for drug discovery: methods and applications, *Nat Rev Drug Discov,* 2004, **3**, 935–949. (DOI: 10.1038/nrd1549)

18. M. Karplus, and J. McCammon, Molecular dynamics simulations of biomolecules, *Nat Struct Mol Biol,* 2002, **9**, 646–652. (DOI: 10.1038/nsb0902-646)

19. X. Gonze et al., The Abinit project: Impact, environment and recent developments, *Comput Physics Communications,* 2020, **248**, 107042. (DOI: 10.1016/j.cpc.2019.107042)

20. P. Giannozzi et al., QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, *J. Phys.: Condens. Matter*, 2009, **21**(39), 395502. (DOI: 10.1088/0953-8984/21/39/395502)
21. T. D. Kühne et al., CP2K: An electronic structure and molecular dynamics software package -Quickstep: Efficient and accurate electronic structure calculations, *J. Chem. Phys.*, 2020, **152**, 194103. (DOI: 10.1063/5.0007045)
22. M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, and W. A. de Jong, NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations, *Comput. Phys. Commun.*, 2010, **181**, 1477. (DOI10.1016/j.cpc.2010.04.018)
23. G. M. J. Barca, Recent developments in the general atomic and molecular electronic structure system, *J. Chem. Phys.*, 2020, **152**, 154102. (DOI: 10.1063/5.0005188)
24. A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J Plimpton, LAMMPS - a flexible simulation tool for particle- based materials modeling at the atomic meso, and continuum scales, *Comp Phys Comm*, 2022, **271**,10817. (DOI: 10.1016/j.cpc.2021.108171)
25. A. García, Siesta: Recent developments and applications, *J. Chem. Phys.*, 2020, **152**, 204108. (DOI: 10.1063/5.0005077)
26. U. Alekseeva, G. Michalicek, D. Wortmann, and S. Blügel, Hybrid Parallelization and Performance Optimization of the FLEUR Code: New Possibilities for All-Electron Density Functional Theory, 2018, In: M. Aldinucci, L. Padovani, M. Torquati, (eds) Euro-Par 2018: Parallel Processing. Euro-Par 2018. Lecture Notes in Computer Science(), vol 11014. Springer, Cham. (DOI: 10.1007/978-3-319-96983-1_52)
27. J. Ren, W. Liang, and M. H. Whangbo, Crystal and electronic structure analysis using CAESAR, PrimeColor Software Inc., 1998, Cary, North Carolina, USA.
28. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX*, 2015, **1–2**, 19–25. (DOI: 10.1016/j.softx.2015.06.001)
29. J. C. Phillips et al., Scalable molecular dynamics on CPU and GPU architectures with NAMD, *Journal of Chemical Physics*, 2020, **153**, 044130. (DOI: 10.1063/5.0014475)
30. O. Trott and A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J Comput Chem.*, 2010, **31**(2), 455-61. (DOI: 10.1002/jcc.21334)
31. H. M. Yu and A. S. Banerje, Density functional theory method for twisted geometries with application to torsional deformations in group-IV nanotubes, *Journal of Computational Physics,* 2022, **456**, 111023. (DOI: 10.1016/j.jcp.2022.111023)
32. B. J. Min, H. K. Jeong, and C. Lee, Plane wave density functional theory studies of the structural and the electronic properties of amino acids attached to graphene oxide via peptide bonding, *Journal of the Korean Physical Society*, 2015, **67**, 507–511. (DOI: 10.3938/jkps.67.507)
33. V. Korpelin, T. Kiljunen, M. M. Melander, M. A. Caro, H. H. Kristoffersen, N. Mammen, V. Apaja, and K. Honkala, Addressing Dynamics at Catalytic Heterogeneous Interfaces with DFT-MD: Anomalous Temperature Distributions from Commonly Used Thermostats, *The Journal of Physical Chemistry Letters*, 2022, **13** (11), 2644-2652. (DOI: 10.1021/acs.jpclett.2c00230)
34. J. Lan, V. V. Rybkin, and M. Iannuzzi, Ionization of Water as an Effect of Quantum Delocalization at Aqueous Electrode Interfaces, *The Journal of Physical Chemistry Letters*, 2020, **11** (9), 3724-3730. (DOI: 10.1021/acs.jpclett.0c01025)
35. S. Aliebrahimi, K. S. Montasser, S. N. Ostad, S. S.Arab, and L. Karami, Identification of Phytochemicals Targeting c-Met Kinase Domain using Consensus Docking and Molecular Dynamics Simulation Studies, *Cell Biochem Biophys.,* 2018, **76**(1-2), 135-145. (DOI: 10.1007/s12013-017-0821-6)

36. R. Majumder and M. Mandal, Screening of plant-based natural compounds as a potential COVID-19 main protease inhibitor: an in silico docking and molecular dynamics simulation approach, *Journal of Biomolecular Structure and Dynamics*, 2022, **40**, 696-711. (DOI: 10.1080/07391102.2020.1817787)

37. A. Rabia, F. Tumino, A. Milani, V. Russo, A. L. Bassi, S. Achilli, G. Fratesi, G. Onida, N. Manini, Q. Sun, W. Xu, and C. S. Casari, Scanning tunneling microscopy and Raman spectroscopy of polymeric sp–sp^2 carbon atomic wires synthesized on the Au(111) surface, *Nanoscale*, 2019, **11**, 18191-18200. (DOI: 10.1039/C9NR06552K)

38. C. Tantardini, A. G. Kvashnin, and D. Ceresoli, GIPAW Pseudopotentials of d Elements for Solid-State NMR, *Materials* (Basel), 2022, **15**(9), 3347. (DOI: 10.3390/ma15093347)

39. A. H. Mazurek, Ł. Szeleszczuk, and D. M. Pisklak, A Review on Combination of Ab Initio Molecular Dynamics and NMR Parameters Calculations, *Int J Mol Sci.*, 2021, **22**(9), 4378. (DOI: 10.3390/ijms22094378)

40. I. Zeydabadi-Nejad, N. Zolfaghari, M. M. Mashhadi, M Baghani, and M. Baniassadi, Anatase TiO_2 nanotubes as Li-ion battery anodes: a molecular dynamics study of Li-ion adsorption on anatase nanotubes, *Sustainable Energy Technologies And Assessments*, 2021, **47**, 101438. (DOI: 10.1016/j.seta.2021.101438)

41. M. Aslan, J. B. A. Davis, and R. L. Johnston, Global Optimization of Small Bimetallic Pd-Co Binary Nanoalloy Clusters: A Genetic Algorithm Approach at the DFT Level, *Physical Chemistry Chemical Physics*, 2016, **18**, 6676-6682. (DOI: 18:6676-6682.10.1039/C6CP00342G)

42. J. Evtimova, E. Drioli, and G. De Luca, A Density Functional Theory Study of Hydrogen Occupation in VNiTi Alloys Used for Dense Metal Membranes, *Journal of Alloys and Compounds*, 2016, **665**, 225-230. (DOI: 10.1016/j.jallcom.2016.01.037)

43. L. Katz and H. D. Megaw, The structure of potassium niobate at room temperature: the solution of a pseudosymmetric structure by Fourier methods, *Acta Crystallographica*, 1967, **22**, 639. (DOI: NA)

44. P. Blaha, K. Schwarz, F. Tran, R. Laskowski, G. K. H. Madsen, and L. D. Marks, WIEN2k: An APW+lo program for calculating the properties of solids., *J. Chem. Phys*., 2020, **152**, 074101. (DOI: 10.1063/1.5143061)

45. M. Profeta, F. Mauri, and C. J. Pickard, Accurate First Principles Prediction of ^{17}O NMR Parameters in SiO2: Assignment of the Zeolite Ferrierite Spectrum, *J. Am. Chem. Soc.*, 2003, **125**, 541. (DOI: 10.1021/ja027124r)

46. J. V. Hanna, K. J. Pike, T. Charpentier, T. F. Kemp, M. E. Smith, B. E. G. Lucier, R. W. Schurko, and L. S. Cahill, A ^{93}Nb Solid□State NMR and Density Functional Theory Study of Four- and Six-Coordinate Niobate Systems, *Chem. Eur. J.*, 2010, **16**, 3222. (DOI: 10.1002/chem.200901581)

47. github. (2020). GitHub. Retrieved from https://github.com/

48. G. Kresse and J. Hafner, Ab initio molecular dynamics for liquid metals, *Phys. Rev. B*, 1993, **47**, 558(R). (DOI: 10.1103/PhysRevB.47.558)

49. C. Tantardini, A. G. Kvashnin, and D. Ceresoli, GIPAW Pseudopotentials of d Elements for Solid-State NMR, *Materials* 2022, **15**, 3347. (DOI: 10.3390/ma15093347)

50. R. Rozza and F. Ferrante, Computational study of water adsorption on halloysite nanotube in different pH environments, *Applied Clay Science*, 2020, **190**, 105589. (DOI: 10.1016/j.clay.2020.105589)

51. C. Cheng, W. Song, Q. Zhao, and H. Zhang, Halloysite nanotubes in polymer science: purification, characterization, modification and applications, *Nanotechnology Reviews*, 2020, **9**, 323-344. (DOI: 10.1515/ntrev-2020-0024)

52. S. Satish, M. Tharmavaram, and D. Rawtani, Halloysite nanotubes as a nature's boon for biomedical applications, *Nanobiomedicine (Rij)*, 2019, **6**, 1849543519863625. (DOI: 10.1177/1849543519863625)