DIADEM - Bourse de Master - 2026/2027

Machine learning assisted crystal structure prediction of transition-metal hemicarbides (M_2C)

PI: Dr. Santanu Saha (CPJ), Dr. Sylvian Cadars (CNRS CR)
University of Limoges, CNRS, IRCER, UMR 7315, F-87000, Limoges, France
Email: santanu.saha@unilim.fr, sylvian.cadars@unilim.fr

Collaboration: Prof. Dr. Stefan Goedecker
Dept. of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland
Email: stefan.goedecker@unibas.ch

Scientific Description

Transition-metal carbides have attracted renewed interest owing to their exceptional properties with broad potential applications, spanning refractory and hard materials, superconductivity and HER/OER electrocatalysis [1]. These materials uniquely combine ceramic hardness, thermal and chemical stability with metallic conductivity. The origin of their multifunctional properties can be traced down to the intricate interplay between metallic and covalent bonding, which governs their structural, electronic and mechanical behaviour [2, 3].

Among different metal-to-carbon stoichiometries, metal mono-carbides (MC) and metal hemicarbides (M_2C) have been extensively studied in 1970s and 1980s with particular emphasis on understanding of their crystal structure, phase diagram and optimization of synthesis routes [4, 5]. Structurally, these systems can be described as the combination of a metal sublattice forming an underlying core framework with occupation of the interstitial voids by the relatively small C atoms. Experimental reports establish that the metal sublattice in MC compounds primarily adapt FCC-derived arrangement with C atoms filling all the octahedral interstitial sites. The ordered occupation of the octahedral sites gives rise to the characteristic rosk-salt structure, in which the metal and carbon atoms form two interpenetrating FCC sublattices. In case of larger metal ions, the MC compound adapts hexagonal (AA stacking) arrangement [6]. Recently, high entropy carbides stabilizing in rock-salt phase have been reported [7]. Thus, the structural chemistry of MC systems is comparatively well established.

On the contrary, the M_2C systems are characterized by HCP-derived metal sublattice framework with partial occupation of the interstitial voids by the C atoms. The low carbon content with respect to the total number of available interstitial voids naturally leads to additional degrees of freedom associated with different degree of C/vacancy ordering [4, 5]. Furthermore, variations in the size and electronic configuration of the metal atoms, along with the differences in M-C bonding interactions promotes distortions of the metal framework and heterogeneity in local-atomic arrangements. Consequently, this combination of structural features gives rise to multiple competing polymorphs with distinct crystal symmetries and local structural motifs. Despite these interesting aspects, the experimental exploration and characterization of M_2C compounds have been challenging due to the requirement of high temperature and precise control of processing conditions for their synthesis; thereby leading to difficulties in stabilization and characterization of individual polymorphs [8].

A natural pathway towards the exploration of the complex phase-space of M_2C systems and parallely addressing the knowledge gap is the adaptation of suitable ab-initio methods (primarily based on density functional theory) such as crystal structure prediction (CSP). However, their application to M_2C systems

quickly becomes prohitive due to the significant increase in computational costs associated with the requirement of large supercells for an explicit representation of different disorders. In principle, machine learned interatomic potentials (MLIPs) can be utilised with CSP to alleviate these bottlenecks. But, they have seldom been used or tested for these specific complex disordered systems [9]. For instance for Mo_2C, only four experimental phases have been reported, i.e. α-Mo_2C ($Pbcn$), β-Mo_2C ($P63/mmc$), C_6-CdI_2 antisite ($P\bar{3}m1$) and the ε-Fe_2N ($P\bar{3}1m$) [4, 5]. Additionally, few other polymorphs have been proposed theoretically [10, 11, 12]. All these reported structures predominantly comprise of HCP-derived metal frameworks [13]. But, this conventional picture has been recently challenged by our work in Ref. [14] which reports multiple non-HCP derived phases, as shown in Fig.(1). With the structural flexibility to accommodate different disorders, the M_2C systems are strong contenders for the experimental realization of high entropy carbides considering the recent success of MC systems [7].

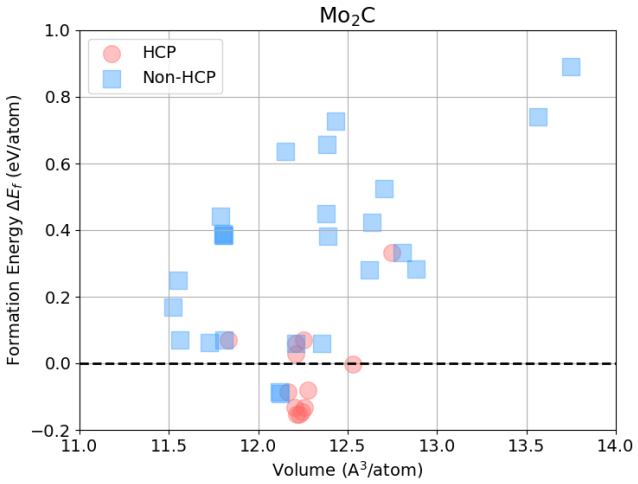


Figure 1: Volume ($\AA^3$/atom) vs Formation energy (eV/atom) of explored Mo_2C phases following Ref. [14] . Here, HCP represents the Mo_2C polymorphs with hexagonal close packed lattice. Non-HCP represents the remaining polymorphs which cannot be structurally represented via the HCP lattice arrangement.

Within this master's thesis project, we aim to explore the phase space of 2-3 representative M_2C compositions using the Minima-Hopping (MH) crystal structure prediction method [15, 16], which has the capability to efficiently explore complex chemical systems [17, 18, 19]. The MH runs will be carried out with universal MACE foundation models [20, 21], to describe the atomic interactions. The primary goal is to understand and map structural diversity across different chemistries. In order to capture different metal-carbon interactions across the transition metal series, representative M_2C compositions will be selected from the $3d$, $4d$, and $5d$ transition-metal chemistries. The generated structures will be filtered, clustered, and analyzed to identify structural motifs and assess their relative thermodynamic stability. This will be subsequently followed by DFT validation primarily through open sourced Quantum-Espresso (QE) package [22]. Where feasible, the resulting DFT dataset will be used to fine-tune the foundation model to improve accuracy of the MLIP and thereby enable more efficient exploration of the selected M_2C compositions.

This work will lead to a better understanding of the diversity of structural motifs across different M_2C chemistries, the role of entropy in phase stabilization and free-energy landscapes, and other thermodynamic properties. In addition, it will provide further insight into the structure-property relationship, particularly with respect to electronic and mechanical properties. The success of this work, both in terms of physico-chemical understanding and computational methodology will form a strong basis for further exploration of multinary M_2C systems.

Scientific Team

The proposed master project will be carried out under the supervision of **Dr. Santanu Saha** (Chair Junior Professor CPJ, PI) and **Dr. Sylvian Cadars** (CNRS CR, co-PI) based at the **Institute of Research for Ceramics, IRCER CNRS UMR 7315 Laboratory and Université de Limoges, 87000 Limoges**. Dr. Saha and Dr. Cadars have extensive experience in application of CSP methods on investigations of different chemical systems [17, 18, 23]. They are co-PIs of the PEPR-DIADEM AMADEUS project (ANR-23-PEXD-0003) on "Machine learning-assisted automated design of high-entropy perovskite-type oxide ceramics for energy applications". In addition, Dr. Saha is also PI of the project on "Ab-initio design of high entropy ceramics" (CPJ fund ANR-22-CPJ1-0039-01). The proposed work will be build upon the computational methodologies and tools developed within these project and tools available in the modelling group based at IRCER, Limoges.

This work will be carried out in collaboration with the team of **Prof. Dr. Stefan Goedecker** based at the **Dept. of Physics, University of Basel, Switzerland**. The minima hopping method was developed by the Goedecker group and has been successfully applied on different class of complex materials using DFT, force-field and recently MLIPs [19, 24, 25]. The Goedecker group will share their expertise in adapting the workflow associated with MH based data generation and their use in fine-tuning of MACE potentials. If required, they will assist in introduction of relevant additional features or developments within MH algorithm for accelerated searches of M_2C systems.

Currently, the scientific team does not have any funding on this scientific topic. However, based on the acceptance and progress of the current project, the involved PIs are planning to submit funding proposals in future regional and national funding (ANR) calls (2027-2028).

References

[1] Y. Xiao et al. *J. Mater. Chem. A*, 4(27):10379–10393, 2016.
[2] VI Ivashchenko et al. *J. Appl. Phys.*, 125(23), 2019.
[3] W. S. Williams. *Mater. Sci. Eng. A.*, 105:1–10, 1988.
[4] J. Dubois et al. *Acta Metall.*, 36(8):1891–1901, 1988.
[5] T. Epicier et al. *Acta Metall.*, 36(8):1903–1921, 1988.
[6] A.S. Kurlov and A.I. Gusev. *Inorganic Materials*, 42(2):121–127, 2006.
[7] P. Sarker et al. *Nat. Commun.*, 9(1):4980, 2018.
[8] B. Deng et al. *Nat. Commun.*, 13(1):262, 2022.
[9] E. A. Bejarano et al. *Phys. Chem. Chem. Phys.*, 27(41):22240–22250, 2025.
[10] W. Sun et al. *Ceram. Inter.*, 50(11):20796–20801, 2024.
[11] J. Lei et al. *J. Mater. Chem. C*, 5(14):3438–3444, 2017.
[12] Y. Guo et al. *Struc. Sci.*, 78(6):848–856, 2022.
[13] C. Deng et al. *Mater. Res. Exp.*, 11(5):056512, 2024.
[14] Asif Iqbal et al. Polymorphism in di-molybdenum carbides mo_2c. *(to be submitted)*, 2026.
[15] S. Goedecker. *J. Chem. Phys.*, 120(21):9911–9917, 2004.
[16] M. Krummenacher et al. *SoftwareX*, 25:101632, 2024.
[17] S. Saha et al. *Phys. Rev. B*, 102(2):024519, 2020.
[18] S. Saha et al. *Phys. Rev. Materials*, 7(5):054806, 2023.
[19] M. Krummenacher et al. *arXiv preprint arXiv:2501.18372*, 2025.
[20] I. Batatia et al. *Adv. Neural Inf. Process. Syst.*, 36:55260–55284, 2023.
[21] I. Batatia et al. *arXiv preprint arXiv:2510.25380*, 2025.
[22] P. Giannozzi et al. *J. Phys. Cond. Matt.*, 29(46):465901, 2017.
[23] S. Cadars et al. *Phys. Chem. Chem. Phys.*, 27(4):2158–2176, 2025.
[24] M. I. Safa et al. *Nanoscale Adv.*, 2026.
[25] O. Tayfuroglu et al. *Commun. Phys.*, 2026.

Computational Methods

Crystal Structure Prediction (CSP): The phase space exploration of 2-3 representative M_2C systems will be carried out using the Minima Hopping (MH) method [1, 2]. It has been successfully applied on different class of complex materials using DFT, force-field and recently MLIPs spanning clusters [3], molecular crystals [4], high pressure ternay hydrides [5] and metal oxides [6]. The established MLIP driven MH workflow will form the basis for proposed studies in this project.

Density Functional Theory (DFT) Calculations: DFT calculations will be performed using the open source Quantum-Expresso software package [7]. Wherever feasible, the DFT dataset will be used for fine-tuning the foundation model to improve accuracy of the MLIP and thereby enable more efficient exploration of the selected M_2C compositions. In addition, DFT will also be used for investigation of low-energy polymorphs with focus on their electronic, charge-density and mechanical properties.

Machine Learned Interatomic Potentials (MLIP): The CSP exploration of the selected 2-3 M_2C compositions will be carried out using the MACE foundation models [8, 9]. Post screening of the generated structures, the final dataset will be evaluated using DFT calculations and used for fine-tuning of the MLIPs.

MLIP assissted CSP workflow: The initial testing and adaptation of the MH runs will be carried out on one of the selected M_2C composition, where the interatomic interactions will be described by the MACE foundation potentials (MH-MACE workflow). The simulation via MH-MACE workflow will lead to generation of a large number of structures. Post screening of the duplicate structures within the initial generated dataset, DFT calculations will be carried out on these pruned dataset for fine-tuning/validation of the MACE-MH1 potentials. In the next step, the combination of fine-tuned MACE with MH will be used to improve upon the exploration of the configuration space to identify low energy structures with different structural motifs. Once established, the workflow will be applied on other representative M_2C systems.

Phase-space exploration of the M_2C Compositions: The dataset of structures generated using the MH-MACE workflow will be first cleaned of duplicate structures using appropriate structural fingerprint descriptors [10]. The cleaned dataset will then be analysed thoroughly to identify structural trends, local atomic arrangements and important structural motifs among different representative compositions. A small selected set of low energy polymorphs will be further evaluated using DFT to investigate their electronic, charge density and mechanical properties. The structural analysis will be carried out using the set of structural analysis tools developed within our group, which are based on atomic simulation environment (ASE) [11] and Pymatgen [12] packages.

References

[1] S. Goedecker. *J. Chem. Phys.*, 120(21):9911–9917, 2004.
[2] M. Krummenacher et al. *SoftwareX*, 25:101632, 2024.
[3] M. I. Safa et al. *Nanoscale Adv.*, 2026.
[4] M. Krummenacher et al. *arXiv preprint arXiv:2501.18372*, 2025.
[5] M. Gubler et al. *New J. Phys.*, 25(12):123008, 2023.
[6] E. K. Rahmatizad et al. *J. Comp. Chem.*, 42(10):699–705, 2021.
[7] P. Giannozzi et al. *J. Phys. Cond. Matt.*, 29(46):465901, 2017.
[8] I. Batatia et al. *Adv. Neural Inf. Process. Syst.*, 36:55260–55284, 2023.
[9] I. Batatia et al. *arXiv preprint arXiv:2510.25380*, 2025.
[10] L. Himanen et al. *Comp. Phys. Commun.*, 247:106949, 2020.
[11] A. H. Larsen et al. *J. Phys. Cond. Matt.*, 29(27):273002, 2017.
[12] S. P. Ong et al. *Comp. Mater. Sci.*, 68:314–319, 2013.

Departement Physik
Universität Basel, Klingelbergstrasse 82
CH-4056 Basel

Prof. Dr. Stefan Goedecker
Tel. +41-61-267-3743
Fax +41-61-267-1349
stefan.goedecker@unibas.ch

Basel, 1. September 2026

Letter of support for the project M̈achine learning assisted crystal structure prediction for phase space exploration of transition-metal hemicarbides

Machine learned potentials are at present revolutionizing atomistic simulations in materials science, physics and chemistry. They allow us to calculate energy and forces at a small fraction of the cost of a density functional calculation without significantly compromising its accuracy. In this way systems of unprecedented size and complexity can now be studied. The project of Dr. Santanu Saha is a pioneering one in this context. By studying large cells of hemicarbides with structure prediction methods, it is to be expected that new structural motifs with interesting short and long range order will be found. Since the structure determines all properties and functionalities of materials, this will lead to the discovery of novel properties. In particular, the importance and effects of entropy on phase diagrams and other thermodynamic properties will be of interest.
The computational physics group at the University of Basel will be glad to collaborate on this project with Santanu. In the group we have extensive experience with structure predictions using machine learned potentials. From our side, it would be highly interesting to apply our simulation methods to this very important class of transition-metal carbides.

Prof. Dr. Stefan Goedecker