

Б. СЕМЕНЮК,<sup>1</sup> О. ФЕЯ<sup>1,2,3,4</sup>

<sup>1</sup> Київський Академічний Університет

(Бул. Академіка Вернадського, 36, Київ 03142; e-mail: feia@kau.edu.ua)

<sup>2</sup> Інститут металофізики імені Г.В. Курдюмова НАН України

(Бул. Академіка Вернадського, 36, Київ 03142)

<sup>3</sup> Leibniz-Institut für Festkörper- und Werkstoffforschung Dresden e.V.

(Helmholtzstraße 20, Dresden 01069, Deutschland)

<sup>4</sup> Державний університет “Київський авіаційний інститут”

(Просп. Любомира Гузара, 1, Київ 03058)

## ПРОГНОЗУВАННЯ КРИСТАЛІЧНОЇ СТРУКТУРИ З ВИКОРИСТАННЯМ ГЕНЕТИЧНИХ АЛГОРИТМІВ БІБЛІОТЕКИ ASE НА МОВІ PYTHON

УДК 539

*Робота присвячена розробці та впровадженню методики прогнозування кристалічної структури з використанням генетичних алгоритмів, інтегрованих у бібліотеку ASE на мові Python. Поєднуючи гнучкість ASE зі швидкістю класичних релаксаторів і точністю методів DFT, наш підхід значно знижує обчислювальні витрати, водночас зберігаючи прогностичну надійність. Методика була перевірена на поліморфах кремнезему (SiO<sub>2</sub>), де наша система дала змогу успішно обчислити як глобальні, так і локальні мінімуми енергетичного ландшафту. Також досліджено інтеграцію релаксаторів на основі нейронних мереж, таких як MACE та AIMNet2, для подальшого пришвидшення процесу пошуку. Це дослідження закладає основу для ефективного, масштабованого й точного прогностичного моделювання кристалічних матеріалів.*

*Ключові слова:* прогнозування кристалічної структури, генетичний алгоритм, енергетичний ландшафт, поліморфи кристалічного кремнезему, SiO<sub>2</sub>, релаксація структури, ASE, GULP.

### 1. Вступ

Безупинний розвиток людської цивілізації потребує інноваційних технологій. Для їхнього успішного впровадження потрібно створення новітніх матеріалів. Однак прогрес у цій галузі часто уповільнюється традиційними методами дослідження матеріалів, які є фінансово витратними, трудоміс-

ткими й випробовують наполегливість дослідників. Технологічний прогрес уможливив автоматизацію цього процесу шляхом інтеграції нових компонентів, таких як комп'ютерне моделювання й прогнозування структури [1, 2], у стандартний робочий процес дослідження матеріалів.

Основним компонентом цього циклу є технологія прогнозування й синтезу структури. Проблеми пов'язані не лише зі швидкістю та точністю алгоритмів прогнозування, а й із практичною значущістю прогнозованих структур. В ідеалі, такі прогнози повинні сприяти безпосередній експериментальній реалізації запропонованих сполук у лабораторних умовах [2]. Галузь обчислювального прогнозування матеріалів швидко розвивається й швидко опановує нові методи, такі як засоби ма-

---

Цитування: Семенюк Б., Фей О. Прогнозування кристалічної структури з використанням генетичних алгоритмів бібліотеки ASE на мові Python. *Укр. фіз. журн.* **71**, № 6, 558 (2026).

© Видавець ВД “Академперіодика” НАН України, 2026. Стаття опублікована за умовами відкритого доступу за ліцензією CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

шинного навчання, нейронні мережі тощо [3, 4]. Більшість алгоритмів для досягнення цієї мети базуються на знаходженні глобальних і локальних мінімумів багатовимірного енергетичного ландшафту (рис. 1) [5].

Енергетичний ландшафт можна уявити як безперервну поверхню, що складається з нескінченної кількості точок, кожна з яких відповідає певному значенню енергії та, в свою чергу, унікальній атомній конфігурації. У цьому представленні вертикальна вісь відображає енергію системи, тоді як горизонтальні осі відповідають структурним параметрам, які можуть змінюватися (рис. 1). Наприклад, ці параметри можуть бути векторами ґратки й атомними положеннями в елементарній комірці, чия колективна зміна визначає доступні області ландшафту.

У контексті прогнозування структури центральною метою є визначення мінімумів енергії, як глобальних, так і локальних, на потенціальному енергетичному ландшафті. Глобальний мінімум відповідає найбільш термодинамічно стабільній структурі, яка також є найімовірнішим кандидатом для успішного експериментального синтезу. Водночас, локальні мінімуми представляють метастабільні конфігурації, які, хоча й менш стабільні, все ж можуть виявлятися експериментально за певних термодинамічних або кінетичних умов.

Для ідентифікації цих мінімумів використовуються різноманітні пошукові алгоритми, такі як перестрибування через басейн (basin hopping) [6], стрибки по мінімумах (minima hopping) [7], метадинаміка [8], імітація відпалу (simulated annealing) [9], а також оптимізація рою частинок (particle swarm optimization) [10]; див. рис. 2, *a–2, e*). У цій роботі ми використовуємо генетичний алгоритм (рис. 2, *f*) [11], який дає можливість ефективно досліджувати й аналізувати великі області енергетичного ландшафту з прийнятною точністю.

Метою цієї роботи є розробка ефективної методології дослідження енергетичних ландшафтів шляхом поєднання генетичних алгоритмів, реалізованих у бібліотеці Python ASE (Atomic Simulation Environment) [12], з обчислювальними пакетами GULP [13] і Quantum ESPRESSO [14]. Цей інтегрований підхід був розроблений для забезпечення точного й прийнятного з погляду обчислювальних ресурсів способу визначення глобальних мінімумів складних енергетичних поверхонь.

## 2. Методологія обчислень

Успішне застосування генетичних алгоритмів залежить від використання обчислювальних “калькуляторів”. Ці інструменти повинні бути здатні не лише оцінювати повну енергію структур-кандидатів, але й виконувати структурні релаксації, тобто оптимізувати атомні конфігурації, щоб знайти найближчий локальний мінімум енергії на ландшафті.

Генетичний алгоритм реалізовано нами засобами бібліотеки ASE, створеної на основі мови Python [12], тоді як структурні релаксації виконуються за допомогою двох додаткових калькуляторів, GULP [13] і Quantum ESPRESSO [14]. GULP використовує класичні потенціали на основі закону Кулона [13], що призводить до нижчої точності, але дає змогу проводити оцінювання набагато швидше. Натомість Quantum ESPRESSO спирається на теорію функціоналу густини (DFT) [15, 16], забезпечуючи значно вищу точність завдяки значно більшим обчислювальним витратам. Використовуючи сильні сторони обох інструментів, наш підхід передбачає використання GULP для швидкого початкового скринінгу структур-кандидатів, із подальшим уточненням за допомогою Quantum ESPRESSO для забезпечення надійного й точного визначення мінімумів енергії.

Генетичний алгоритм працює як ітераційний процес [12] (рис. 3), де кожна ітерація називається поколінням (або популяцією). Пошук починається з початкової популяції випадково згенерованих структур. Після оцінки їхніх енергій отриманий розподіл відповідає набору випадково розкиданих точок по всьому енергетичному ландшафту.

На наступному кроці структури-кандидати оцінюються й впорядковуються відповідно до їхніх

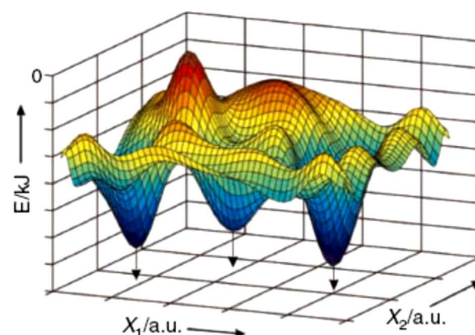


Рис. 1. Багатовимірні енергетичні ландшафти [5]

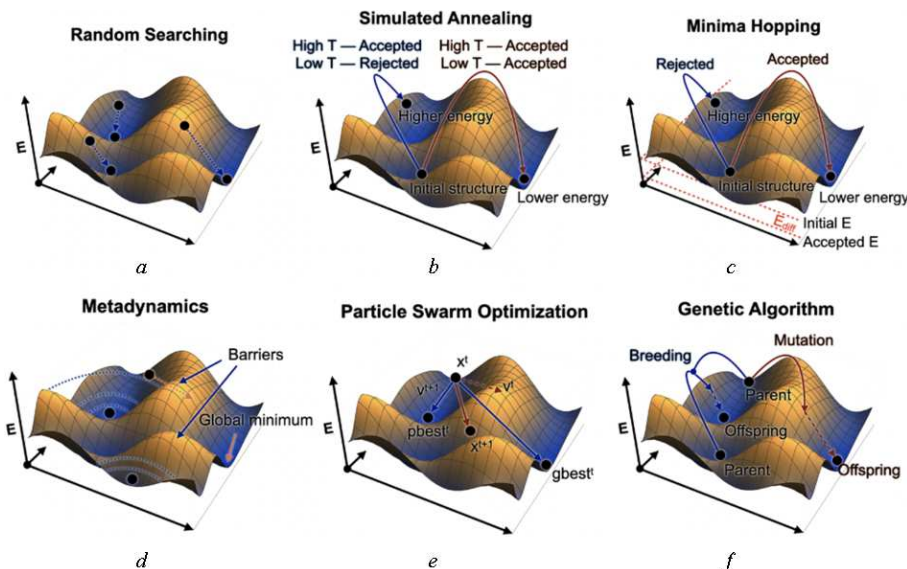


Рис. 2. Різні методи оптимізації глобальних мінімумів енергетичного ландшафту [11]

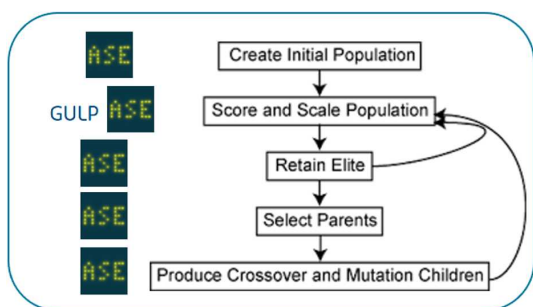


Рис. 3. Ітеративний генетичний алгоритм [11, 12]

енергій. Кожна структура спочатку релаксується, а потім визначається її кінцева енергія. Структури впорядковуються за цими значеннями, причому конфігурація з найнижчою енергією представляє найстабільніше розташування. Наступне покоління будується з суміші вибраних фракцій: це – підмножина найкращих структур, перенесених з попереднього покоління; частина мutowаних варіантів, отриманих шляхом застосування операторів мутації до наявних структур, і частка новостворених випадкових структур, аналогічних тим, що належать початковій популяції. Цей цикл релаксація–оцінка–вибір ітеративно повторюється, поступово спрямовуючи пошук областей ландшафту з нижчою енергією.

Оператори мутації відіграють вирішальну роль у цій еволюційній схемі, даючи змогу досліджува-

ти локальні мінімуми навколо батьківських структур. Завдяки контрольованим збуренням вони запобігають передчасній конвергенції, даючи алгоритму можливість уникнути локальних пасток та ефективніше вибірково досліджувати навколишні області енергетичного ландшафту. Нами було використано три типи операторів мутацій. Оператор *спарювання* (*cut-and-slice*) поєднує дві батьківські структури, розрізаючи їх вздовж площини й обмінюючи їх частини для створення потомства, тим самим уможливаючи перенесення стабільних атомних розташувань від однієї структури до наступного покоління. Оператор *деформування* (*strain*) вводить випадкові зміни до параметрів елементарної комірки, полегшуючи дослідження сусідніх областей енергетичного ландшафту. Нарешті, оператор *м'якої мутації* (*softmutation*) збурює атомні положення вздовж низькочастотних коливальних мод, оцінених за допомогою простої моделі гармонічних зв'язків, тим самим сприяючи ефективній вибірці структурно релевантних конфігурацій [17–20].

Під час виконання алгоритму, калькулятор GULP в основному використовувався для проведення швидких оцінок у кількох поколіннях (популяціях). Пошук завершувався після досягнення заздалегідь визначеного критерію збіжності, який визначався як повторна поява ідентичних структур на діаграмі енергія–придатність, що буде де-

тальніше розглянуто нижче. Ця процедура дає набір структур-кандидатів із симетрією та геометрією, що дуже нагадують цільову (еталонну) фазу. Ці вибрані структури згодом уточнюються шляхом остаточної релаксації за допомогою калькулятора Quantum ESPRESSO. Важливо, що цей етап вимагає лише кількох десятків циклів обчислень, на відміну від сотень циклів, виконаних на етапі початкового дослідження за допомогою GULP, завдяки чому можна досягти значної економії загальних обчислювальних витрат.

Виконуючи розрахунки методом DFT, як початкову інформацію ми використовували кристалічні структури, попередньо релаксовані за допомогою GULP. Структурні оптимізації проводилися за допомогою Quantum ESPRESSO v6.8 [14, 21, 22] з використанням квазіньютонівського алгоритму BFGS, мінімізуючи як загальну енергію, так і міжатомні сили. Обмінно-кореляційний функціонал було розглянуто в рамках узагальненого градієнтного наближення (generalized gradient approximation GGA) Perdew–Burke–Ernzerhof (PBE) [23], реалізованого з використанням псевдопотенціалів PBEsol і метода проективно-доповнених плоских хвиль [24]. Використана енергія відсічення плоскої хвилі, що дорівнює 80 Ry. Релаксації та розрахунки самоузгодженого поля проводилися за допомогою схеми Монкхорста–Пака з використанням сітки розміром  $8 \times 8 \times 8$  з  $k$ -точок для моделювання зони Бріллюена.

### 3. Результати та обговорення

Для перевірки й підтвердження ефективності нашої методики ми обрали матеріал, який виразно демонструє наявність глобального й багатьох локальних мінімумів в енергетичному ландшафті. Діоксид кремнію ( $\text{SiO}_2$ ) [25] було обрано для цієї мети завдяки його багатому поліморфізму (рис. 4). Зокрема, стабільна фаза за умов навколишнього середовища ( $\alpha$ -кварц) служить глобальним мінімумом, тоді як метастабільні фази (такі як  $\beta$ -кварц,  $\alpha$ - та  $\beta$ -кристобаліт, тридиміт, кеатит і стішовіт) є локальними мінімумами.

Результати, отримані для послідовних популяцій, зручно візуалізувати за допомогою діаграм енергія–придатність (рис. 5), де енергія кожної структури відображається залежно від номера її покоління та її позиції в цьому поколінні. На цих діаграмах кожна структура представлена кольо-

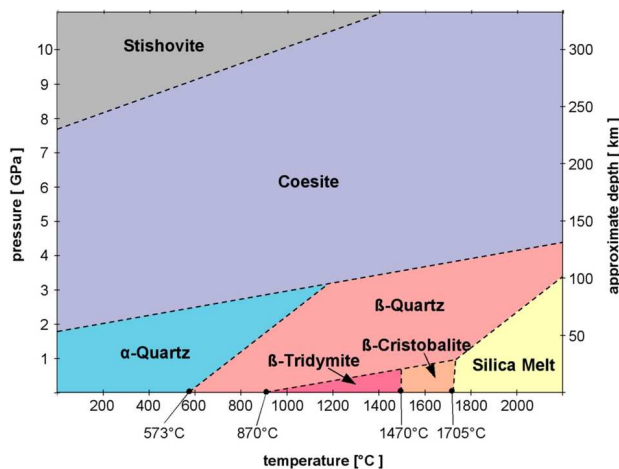
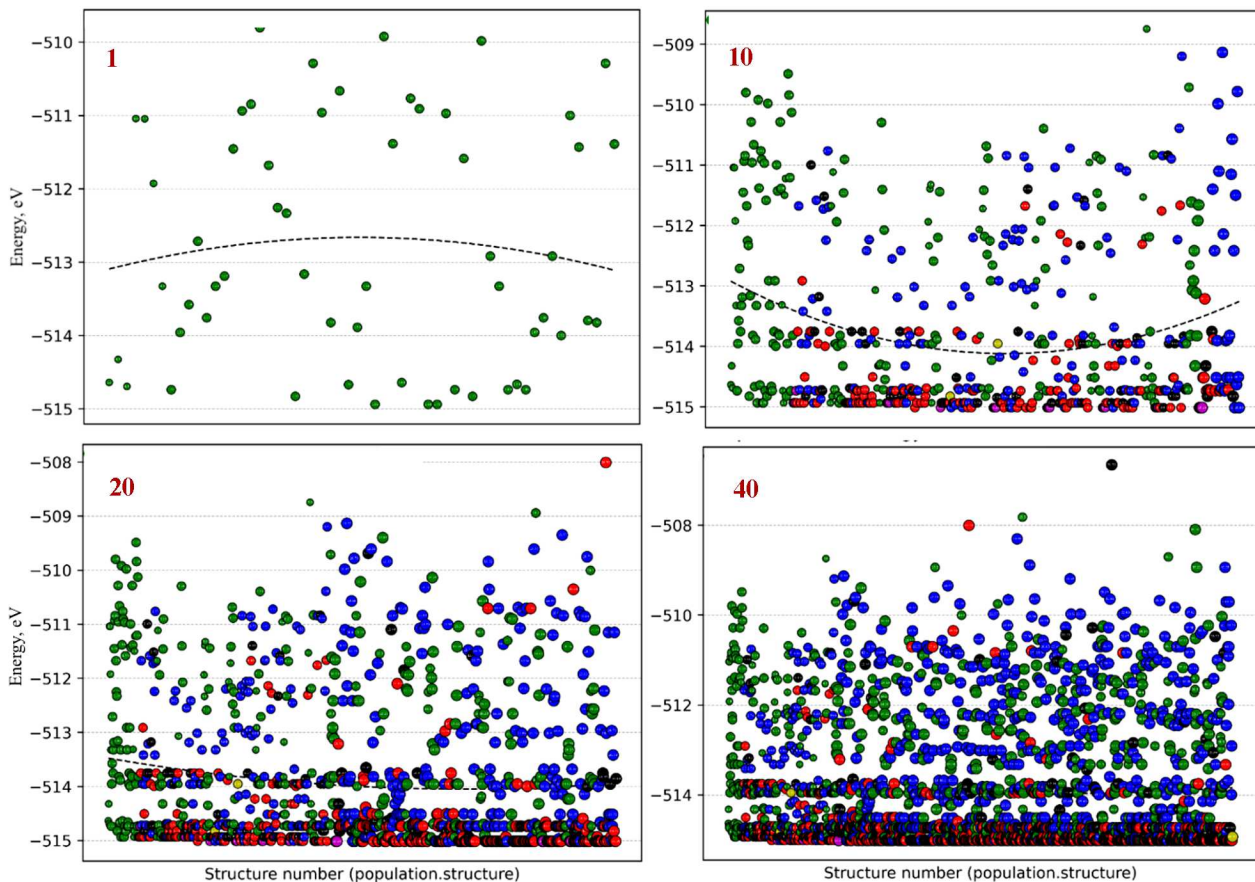


Рис. 4. Фазова діаграма кремнезему [25]

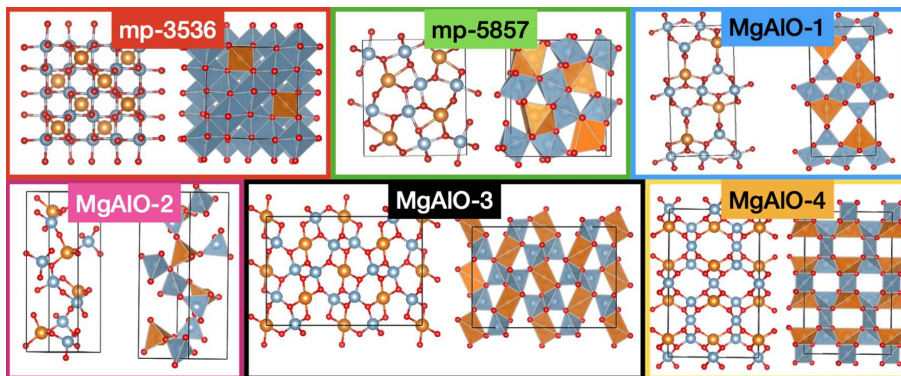
ровою крапкою: початкова випадково згенерована популяція відображається зеленим кольором; структури з найкращими показниками виділені фіолетовим, синім, жовтим і червоним кольорами (що відповідає їхньому створенню різними операторами мутації); а чорні крапки позначають структури, що були відкинуті через надто великий об'єм їхньої елементарної комірки.

Порівняння всіх чотирьох діаграм демонструє виразну закономірність: це – наявність чітких горизонтальних ліній, утворених структурами з різних поколінь, кожна з яких розташована на характерному енергетичному рівні. Ці лінії відповідають локальним і глобальним мінімумам енергетичного ландшафту. У цьому випадку найнижча лінія відповідає  $\alpha$ -кварцу, тоді як вище розташовані лінії відповідають метастабільним фазам, таким як  $\beta$ -кварц,  $\alpha$ - та  $\beta$ -кристобаліт, тридиміт, кеатит і стішовіт. Параметри всіх цих структур були успішно розраховані, за винятком коеситу, який не був ідентифікований через відносно великий розмір його елементарної комірки, оскільки такі масштабні розрахунки не передбачалися в цьому дослідженні. Цей результат підкреслює здатність запропонованої методики враховувати не лише термодинамічно стабільну фазу, а й широкий спектр метастабільних поліморфів, тим самим підкреслюючи її ефективність для прогнозування структур у складних енергетичних ландшафтах.

Для чіткого порівняння передбачені структури разом з відповідними еталонними структурами з бази даних Materials Project [26] представле-



**Рис. 5.** Результати генерації популяції кремнезему. Зелені кола представляють випадково створені структури; рожеві – це найкращі структури з попередньої популяції; сині, жовті й червоні – це структури, отримані варіаційними операторами спарювання, м'якої мутації та деформації, відповідно; чорні – це структури, видалені з пошуку через неадекватні параметри. Енергія вимірюється в електрон-вольтах



**Рис. 6.** Кристалічні структури, що відповідають табл. 2. Помаранчеві сфери представляють атоми Mg, сині сфери відповідають атомам Al, а червоні сфери позначають атоми O. Крім того, помаранчеві багатогранники намальовано навколо позицій Mg, а сині – навколо позицій Al. Усі візуалізації кристалічних структур створено за допомогою програми VESTA [27]

но в табл. 1. Для забезпечення узгодженості кожна структура була ретельно повторно релаксована за допомогою Quantum ESPRESSO з викори-

станням ідентичних обчислювальних параметрів. Структурна подібність потім була кількісно визначена шляхом обчислення косинусної відстані між

Таблиця 1. Результати, отримані еволюційним пошуком (верхній рядок) і релаксацією структур з Materials Project (нижній рядок)

Назва	$abc, \text{Å}$	$\alpha \beta \gamma, ^\circ$	Просторова група	Енергія/атом, Ry	Косинусна відстань
$\alpha$ -кварц	4,961 4,961 5,452	90 90 120	152	-42,767192	0,0
mp-7000	4,96 4,96 5,452	90 90 120	152	-42,767615	
$\beta$ -кварц	5,081 5,081 5,562	90 90 120	180	-42,76719	0,0039
mp-6922	5,081 5,081 5,562	90 90 120	180	-42,767187	
$\beta$ -тридиміт	5,254 5,254 8,575	90 90 120	194	-42,766982	0,0
mp-559091	5,254 5,254 8,575	90 90 120	194	-42,766976	
$\beta$ -кристобаліт	7,430 7,430 7,430	90 90 90	227	-42,766908	0,0355
mp-8352	7,430 7,430 7,430	90 90 90	227	-42,766929	
Стішовіт	4,192 4,192 2,6805	90 90 90	136	-42,762639	0,0069
mp-6947	4,193 4,193 2,680	90 90 90	136	-42,762690	
Флюорит	4,543 4,543 4,543	90 90 90	225	-42,684103	0,02
mp-10064	4,545 4,545 4,545	90 90 90	225	-42,684113	
Гідрофіліт	4,082 5,039 4,497	90 90 120	60	-42,581702	0,0025
mp-10498	4,082 5,041 4,497	90 90 120	60	-42,758837	

Таблиця 2. Система Mg–Al–O. Структури з Materials Project (перші два рядки) і результати еволюційного пошуку (позначені як MgAlO–N)

Назва	$abc (\text{Å})$	$\alpha \beta \gamma (^\circ)$	Просторова група	Енергія/атом, Ry
mp-3536	8,109 8,109 8,109	90 90 90	227	-54,741383
mp-5857	8,672 2,794 9,969	90 90 90	62	-54,735592
MgAlO-1	3,010 14,010 7,006	90 90 90	63	-54,735508
MgAlO-2	4,903 15,213 5,291	90 111,16 90	9	-54,729543
MgAlO-3	14,769 10,307 5,041	90 90 90	43	-54,729471
MgAlO-4	15,922 5,064 9,756	90 90 90	22	-54,726328

fingerprint векторами [28], де значення нуль означало ідеальну відповідність. На основі цього аналізу було підтверджено, що структури, згенеровані нашим методом, не відрізняються від відомих еталонних фаз.

З метою додаткової перевірки ми дослідили систему Mg–Al–O, зосередившись на шпінелі алюмінату магнію  $\text{Mg}(\text{AlO}_2)_2$ . Ця сполука була обрана через її структурну складність і додаткову проблему, пов'язану з її потрійним складом, що робить прогнозування структури особливо складним. Хоча глобальний мінімум у цьому випадку не вдалося визначити, наш підхід виявив кілька раніше невідомих низькоенергетичних метастабільних структур, що робить цю систему вартим уваги прикладом для обговорення (табл. 2).

База даних Materials Project містить відомості лише про дві експериментально підтвержені структури з таким складом: mp-3536 і mp-5857 (див. табл. 2 і рис. 6). Фаза з найнижчою енергією (mp-3536) кристалізується в просторовій групі  $Fd\bar{3}m(227)$  і складається з тетраедрів  $[\text{MgO}_4]$  та октаедрів  $[\text{AlO}_6]$ . Її стандартна елементарна комірка містить 56 атомів, що відповідає 14 атомам в примітивній комірці. Друга експериментально відома структура (mp-5857) має орторомбічну симетрію  $Pnma$ . Примітно, що одна зі структур, отриманих у нашому пошуку, позначена як MgAlO-1 (табл. 2, рис. 6), майже вироджена за енергією зі структурою mp-3536 (її енергія лише на  $8 \times 10^{-5}$  Ry вища) і кристалізується в просторовій групі  $Cmcm(63)$ . У структурі mp-5857 атоми

Mg координуються з сімома атомами O, тоді як атоми Al мають координаційне число шість. На противагу цьому, в MgAlO-1 атоми Mg мають п'ятикратну координацію, але октаедричні одиниці  $[AlO_6]$  все ще зберігаються. У структурах вищих енергій – MgAlO-2, MgAlO-3 і MgAlO-4 (табл. 2, рис. 6) – таких октаедричних будівельних блоків немає; натомість ці фази складаються лише з тетраедрів  $[MgO_4]$  і  $[AlO_4]$ . Це свідчить про те, що для  $Mg(AlO_2)_2$  співіснування тетраедрів, центрованих на Mg і Al, є енергетично не вигідним, тоді як наявність октаедрів  $[AlO_6]$  значною мірою сприяє стабілізації структури, про що свідчать як експериментально підтверджені фази, так і наше передбачення MgAlO-1.

Існує кілька можливих причин, чому глобальний мінімум системи Mg–Al–O не був досягнутий, незважаючи на ідентифікацію численних низькоенергетичних станів. Найбільш ймовірним поясненням є розмір елементарної комірки: за наявності 56 атомів об'єм пошуку стає надзвичайно великим і його важко вичерпно дослідити. Другим фактором є обмеження калькулятора GULP, який спирається на класичні кулонівські взаємодії й тому може не враховувати весь спектр фізичних ефектів, необхідних для точного опису цієї системи. Це свідчить про те, що для покращення пошуку справжнього глобального мінімуму потрібен більш досконалий калькулятор зі збереженням обчислювальної ефективності.

#### 4. Висновки та майбутні вдосконалення

Ми впровадили методику прогнозування кристалічних структур шляхом інтеграції мови Python з релаксаторами GULP і Quantum ESPRESSO. Цей підхід був успішно перевірений на бінарних хімічних системах, але його продуктивність є обмеженою для складніших трійних систем, таких як Mg–Al–O. Використання двох релаксаторів суттєво знизило загальні обчислювальні витрати й виявилось ефективним для визначення як глобальних, так і локальних мінімумів енергії. Водночас це привело до появи нових викликів, зокрема необхідності генерувати вхідні файли для GULP, що знижує загальну зручність використання робочого процесу.

Хоча GULP забезпечує швидкі, але менш точні оцінки, Quantum ESPRESSO, як метод на основі DFT, пропонує набагато вищу точність завдя-

ки значним обчислювальним ресурсам [29]. Нещодавні досягнення в галузі потенціалів нейронних мереж пропонують обладдйливу альтернативу: навчаючись на наборах даних, отриманих методами DFT, вони можуть наблизитися до точності рівня DFT, водночас різко знижуючи обчислювальні вимоги, хоча точність обмежується якістю навчального набору. Щоб подолати ці обмеження, ми врахували псевдопотенціали, отримані на основі машинного навчання, в нашому робочому процесі. Зокрема, ми застосували MACE [30] і AIMNet2 [31] як заміну GULP у рамках генетичного алгоритму. Зараз ми працюємо над оптимізацією їхньої продуктивності під час прогнозування структури.

*Ми вдячні всім нашим колегам за плідні дискусії на семінарських засіданнях КАУ: професору Олександру Кордюку, професору Євгену Леню, доктору Володимирі Беззубі, а також нашому колаборатору з Університету Карнегі-Меллона професору Олександру Ісаєву. О.Ф. дякує Ulrike Nitzsche за технічну допомогу з комп'ютерним кластером IFW Dresden. Ця робота була підтримана Федеральним міністерством досліджень, технологій і космосу Німеччини (BMFTB) в рамках проекту GU-QuMat (01DK24008). Також ця робота підтримується грантом Національної академії наук України для дослідницьких лабораторій молодих вчених "Пошук топологічної надпровідності в купратах та надпровідниках на основі заліза".*

1. D.G. Pettifor. Computer-aided materials design: bridging the gaps between physics, chemistry and engineering. *Phys. Ed.* **32**, 164 (1997).
2. S.G. Louie, Y.H. Chan, F.H. da Jornada, Z. Li, D.Y. Qiu. Discovering and understanding materials through computation. *Nat. Mater.* **20**, 728 (2021).
3. T. Xie, J.C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
4. S. Feng, H. Zhou, H. Dong. Using deep neural network with small dataset to predict material defects. *Mater. Des.* **162**, 300 (2019).
5. M. Jansen. The energy landscape concept and its implications for synthesis planning. *Pure Appl. Chem.* **86**, 883 (2014).
6. D.J. Wales, J.P. Doye. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem. A* **101**, 5111 (1997).

7. B. Schaefer, S. Mohr, M. Amsler, S. Goedecker. Minima hopping guided path search: an efficient method for finding complex chemical reaction pathways. *J. Chem. Phys.* **140**, 214102 (2014).
8. A. Barducci, M. Bonomi, M. Parrinello. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 826 (2011).
9. S. Kirkpatrick, C.D. Gelatt Jr, M.P. Vecchi. Optimization by simulated annealing. *Science* **220**, 671 (1983).
10. Y. Wang, J. Lv, L. Zhu, Y. Ma. CALYPSO: A method for crystal structure prediction. *Comput. Phys. Commun.* **183**, 2063 (2012).
11. Z. Falls, P. Avery, X. Wang, K.P. Hilleke, E. Zurek. The XtalOpt evolutionary algorithm for crystal structure prediction. *J. Phys. Chem. C* **125**, 1601 (2021).
12. A.H. Larsen, J.J. Mortensen, J. Blomqvist, I.E. Castelli, R. Christensen, M. Dulak, J. Friis, M.N. Groves, B. Hammer, C. Hargus, E.D. Hermes, P.C. Jennings, P.B. Jensen, J. Kermode, J.R. Kitchin *et al.* The atomic simulation environment—a Python library for working with atoms. *J. Phys. Condens. Matter.* **29**, 273002 (2017).
13. J.D. Gale, A.L. Rohl. The general utility lattice program (GULP). *Mol. Simul.* **29**, 291 (2013).
14. P. Giannozzi *et al.* QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter.* **21**, 395502 (2009).
15. P. Hohenberg, W. Kohn. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864 (1964).
16. W. Kohn, L.J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133 (1965).
17. D.M. Deaven, K.M. Ho. Molecular geometry optimization with a genetic algorithm. *Phys. Rev. Lett.* **75**, 288 (1995).
18. S. Lysgaard, D.D. Landis, T. Bligaard, T. Vegge. Genetic algorithm procreation operators for alloy nanoparticle catalysts. *Top. Catal.* **57**, 33 (2014).
19. P.B. Jensen, S. Lysgaard, U.J. Quaade, T. Vegge. Designing mixed metal halide ammines for ammonia storage using density functional theory and genetic algorithms. *Phys. Chem. Chem. Phys.* **16**, 19732 (2014).
20. M. Van den Bossche, H. Gronbeck, B. Hammer. Tight-binding approximation-enhanced global optimization. *J. Chem. Theor. Comput.* **14**, 2797 (2018).
21. P. Giannozzi *et al.* Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys. Condens. Matter.* **29**, 465901 (2017).
22. P. Giannozzi *et al.* Quantum ESPRESSO toward the exascale. *J. Chem. Phys.* **152**, 154105 (2020).
23. J.P. Perdew, K. Burke, M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
24. P.E. Blochl. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).
25. K. Royce, C. Baars, H. Viles. Defining damage and susceptibility, with implications for mineral specimens and objects: introducing the mineral susceptibility database. *Stud. Conserv.* **68**, 298 (2023).
26. A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
27. K. Momma, F. Izumi. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. Cryst.* **44**, 1272 (2011).
28. N.E. Zimmermann, A. Jain. Local structure order parameters and site fingerprints for quantification of coordination environment and crystal structure similarity. *RSC Adv.* **10**, 6063 (2020).
29. J.S. Smith, O. Isayev, A.E. Roitberg. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **8**, 3192 (2017).
30. I. Batatia, D.P. Kovacs, G. Simm, C. Ortner, G. Csanyi. MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. *Adv. Neural Inf. Process* **35**, 11423 (2022).
31. R. Zubatyuk, J.S. Smith, J. Leszczynski, O. Isayev. Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network. *Sci. Adv.* **5**, eaav6490 (2019).

Одержано 15.10.25.

Переклад українською О. Войтенка

*B. Semeniuk, O. Feia*

## IMPLEMENTATION OF A METHODOLOGY FOR CRYSTAL STRUCTURE PREDICTION USING GENETIC ALGORITHMS INTEGRATED INTO THE PYTHON ASE LIBRARY

This work is dedicated to the development and implementation of a methodology for crystal structure prediction using genetic algorithms integrated into the Python ASE library. Crystal structure prediction plays a critical role in materials science, chemistry, and nanotechnology, enabling the discovery of novel compounds with tailored properties. By combining the flexibility of ASE with the speed of classical relaxers and the accuracy of DFT-based methods, our approach significantly reduces computational costs while maintaining predictive reliability. The methodology was validated on polymorphs of silica (SiO<sub>2</sub>), where our system successfully recovered both global and local minima of the energy landscape. We also explore the integration of neural network relaxers such as MACE and AIM-Net2 to further accelerate the search process. This study lays the groundwork for efficient, scalable, and accurate predictive modeling of crystalline materials.

*Keywords:* crystal structure prediction, genetic algorithm, energy landscape, crystalline silica polymorphs, SiO<sub>2</sub>, structure relaxation, ASE, GULP.