Aurora Nano Studio

July 1, 2026

Copyright, Proprietary Rights, and Confidentiality Notice

Aurora Nano Studio

Technical, Scientific, and Enterprise Platform Documentation

Certain scientific, mathematical, chemical, physical, computational, architectural, operational, and product-governance descriptions in this document are intentionally presented at a professional validation level rather than as implementation instructions. Aurora Simulations may withhold internal implementation details, optimization strategies, solver integrations, production parameters, infrastructure controls, source code, private kernels, security logic, deployment procedures, and proprietary engineering methods as confidential information and trade secrets.

The information contained in this document is provided for evaluation and informational purposes. It does not constitute a public specification, open-source release, warranty, service-level commitment, professional engineering certification, regulatory submission, or legal guarantee. Platform capabilities, performance, interfaces, deployment targets, validation methods, and product features may evolve over time and may depend on configuration, workload class, deployment environment, customer requirements, and executed commercial agreements.

Any authorized recipient must protect this document using at least the same degree of care used to protect its own confidential technical and commercial information, and in no case less than reasonable care. Unauthorized use, disclosure, reproduction, redistribution, reverse engineering, competitive analysis for replication, or extraction of proprietary methods is strictly prohibited.

Confidential and Proprietary

Aurora Simulations — Aurora Nano Studio

Document Version: 1.0 | Date: July 2026

Aurora Nano Studio

A Unified Scientific Analysis Environment for Atomistic Simulation

White Paper

Organization:	Aurora Simulations
Author:	Dr. Jose Luis Silva
Inquires:	jose.silva@simaurora.com
Nano Studio:	https://simaurora.com
Release v.1.0	01 July 2026

Contents

Abstract 6

1 Introduction 6

2 Structure Representation and Crystallographic Foundations 7

2.1 Lattice Vectors and the Cell Matrix . . . 7

2.2 Symmetry, Space Groups, and Wyckoff Positions . . . 9

2.3 Niggli Reduction . . . 9

2.4 Periodic Boundary Conditions . . . 9

3 Chemical Bonding and Molecular Topology 9

3.1 PBC-Aware Bond Inference . . . 9

3.2 Hydrogen Bond Detection . . . 10

3.3 Effective Coordination Number . . . 10

3.4 Bond-Valence Sums . . . 10

3.5 Madelung Electrostatics . . . 11

3.6 SMILES Ingestion and Molecular Graph . . . 11

3.7 Polymer Chain Construction . . . 11

4 Radial and Angular Distribution Functions 11

4.1 Radial Distribution Function . . . 11

4.2 Angular Distribution Function . . . 12

4.3 Bond-Length Distribution . . . 12

4.4 Static Structure Factor . . . 12

5 Dynamical Analysis and Trajectory Processing 12

5.1 Mean-Square Displacement and Diffusion . . . 12

5.2 Velocity Autocorrelation Function and Vibrational Density of States . . . 13

5.3 Lindemann Melting Index . . . 13

5.4 Trajectory Import, Indexing, and Time Averaging . . . 13

6 Defect Analysis and Local Structure Classification 14

6.1 Centrosymmetry Parameter . . . 14

6.2 Wigner–Seitz Vacancy and Interstitial Mapping . . . 14

6.3 Atomic Strain and Non-Affine Displacement . . . 14

6.4 Common-Neighbor Analysis . . . 15
6.5 Polyhedral Template Matching . . . 15
6.6 Steinhardt Bond-Orientational Order Parameters . . . 15
6.7 Additional Classification Methods . . . 15
6.8 Cluster Analysis . . . 16

7 Volumetric Field Analysis and Electron Density 16
7.1 Supported Scalar Field Types . . . 16
7.2 Derived Fields . . . 16
7.3 Electron Localization Function . . . 17
7.4 Tricubic Interpolation and Upsampling . . . 17
7.5 Marching-Cubes Iso-Surface Extraction . . . 17
7.6 Two-Dimensional Slice Engine . . . 17

8 Surface and Interface Science 18
8.1 Slab Construction from Miller Indices . . . 18
8.2 Lattice Matching for Interfaces . . . 18
8.3 Interface Adhesion Energy . . . 18
8.4 Registry Scans . . . 19
8.5 Adsorption Site Analysis . . . 19
8.6 Computational Input Generation . . . 19

9 Phase Stability and Thermodynamics 19
9.1 Convex Hull Construction . . . 19
9.2 Chemical Potential Diagrams . . . 20

10 Powder Diffraction and Crystallographic Reporting 20
10.1 Powder Diffraction Pattern Simulation . . . 20
10.2 Peak Shape and Pattern Profile . . . 20
10.3 Rietveld Refinement and Related Methods . . . 21
10.4 Crystal Morphology . . . 21
10.5 Patterson Fourier Maps . . . 21

11 Electronic Structure Analysis 22
11.1 Band Structure . . . 22
11.2 Density of States . . . 22

11.3 Brillouin Zone and k-Path Generation . . . 22
11.4 Fermi Surface . . . 22

12 Molecular Simulation Preparation **22**
12.1 Polymer Builder . . . 22
12.2 Simulation Box Packing . . . 22
12.3 Force-Field Assignment . . . 23
12.4 Simulation Deck Generation . . . 23
12.5 Heavy Trajectory Analysis . . . 23

13 Professional Platform Boundary and Scientific Disclosure **23**
13.1 Disclosure Layers . . . 24
13.2 Scientific Contract of a Professional Platform . . . 24
13.3 Professional Language for Proprietary Methods . . . 25

14 Extended Mathematical Derivations for Periodic Geometry **25**
14.1 Metric Tensor and Cell Volume . . . 25
14.2 Reciprocal Lattice and Miller Planes . . . 26
14.3 Fractional Minimum Images in General Cells . . . 26
14.4 Tensor Transformation Between Fractional and Cartesian Frames . . . 26
14.5 Cell Deformation and Finite Strain . . . 27

15 Extended Chemistry and Bonding Derivations **27**
15.1 Bond-Length Residuals and Chemical Plausibility . . . 27
15.2 Bond-Valence Residual and Oxidation-State Consistency . . . 28
15.3 Ionic, Covalent, Metallic, and Hydrogen-Bonding Indicators . . . 28
15.4 Coordination Polyhedra as Local Chemical Units . . . 28
15.5 Crystal-Field and Ligand-Environment Context . . . 29
15.6 Charge, Spin, and Local Moment Consistency . . . 29

16 Statistical Mechanics of Atomistic Observables **30**
16.1 Ensemble Averages and Time Averages . . . 30
16.2 RDF Normalization and Coordination Shell Integrals . . . 30
16.3 Potential of Mean Force . . . 30
16.4 Static Structure Factor and Compressibility Limit . . . 31
16.5 Diffusion Tensor and Anisotropic Transport . . . 31

16.6 Green-Kubo Relations . . . 31
16.7 Uncertainty of Time-Correlation Estimates . . . 32

17 Defect, Strain, and Microstructure Physics Extensions 32
17.1 Least-Squares Origin of the Local Deformation Gradient . . . 32
17.2 Polar Decomposition and Rotation-Strain Separation . . . 32
17.3 Scalar Strain Invariants . . . 33
17.4 Centrosymmetry as a Broken-Inversion Measure . . . 33
17.5 Bond-Orientational Order and Local Symmetry . . . 33
17.6 Misorientation and Grain Context . . . 34

18 Volumetric Field Theory and Electronic Observables 34
18.1 Periodic Grid Integrals . . . 34
18.2 Charge-Difference Density . . . 35
18.3 Gradient, Laplacian, and Critical-Point Context . . . 35
18.4 Electron Localization Function . . . 35
18.5 Poisson Relation and Electrostatic Potential . . . 36
18.6 Isosurfaces as Level Sets . . . 36

19 Diffraction and Reciprocal-Space Reporting Extensions 36
19.1 Scattering Vector and Bragg Condition . . . 36
19.2 Structure Factors and Thermal Displacement . . . 37
19.3 Profile Functions and Residual Metrics . . . 37
19.4 Size-Strain Broadening . . . 37
19.5 Patterson and Pair-Vector Interpretation . . . 38

20 Thermodynamic and Materials-Discovery Derivations 38
20.1 Formation Energy and Reference States . . . 38
20.2 Convex Hull and Energy Above Hull . . . 38
20.3 Chemical-Potential Constraints . . . 39
20.4 Surface and Interface Energies . . . 39
20.5 Adsorption Energy . . . 39
20.6 Grand Potential for Open Systems . . . 39

21 Enterprise Scientific Platform Assurances 40
21.1 From Visualization to Governed Scientific Workflows . . . 40

21.2 Lineage as a Scientific Object 40
21.3 Governance Without Scientific Friction 41
21.4 Compute Routing as a Product Boundary 41
21.5 Confidentiality Statement for Client-Facing Use 41
21.6 Quality Language for Reports 42
21.7 Scientific Review Checklist for Prospective Clients 42

22 Validation and Performance **43**
22.1 Performance Benchmarks 43
22.2 Test Coverage 44
22.3 Stated Limitations 44

23 Applications **44**

24 Conclusion **45**

Abstract

Aurora Nano Studio is a scientific analysis environment for atomistic simulation data, designed to serve researchers in materials science, chemistry, physics, crystallography, and engineering. The system provides an integrated workflow from structure import and construction through a modifier-based analysis pipeline, volumetric field inspection, surface and interface preparation, phase stability analysis, powder diffraction fitting, electronic structure visualization, and molecular simulation preparation. Crystallographic operations are grounded in standard cell-matrix formalism with full support for all seven lattice systems, space-group symmetry, Wyckoff orbit expansion, and Niggli reduction. Bond-order assignments follow established covalent-radii tabulations [4, 5] and extend to hydrogen-bond detection, bond-valence sums [6], and Madelung electrostatics [7]. Dynamical analysis covers the mean-square displacement, velocity autocorrelation function, vibrational density of states, Lindemann melting index, and time-averaged structural properties. Defect and microstructure analysis implements centrosymmetry [19], non-affine strain [20], common-neighbor analysis [21], polyhedral template matching [23], and Steinhardt bond-orientational order [24]. Volumetric analysis supports electron density, electron localization function [30], charge difference density, and spin density through tricubic interpolation and marching-cubes surface extraction [33]. Phase stability is assessed through convex-hull construction and chemical-potential diagrams. Performance benchmarks established on current hardware demonstrate sub-second response for systems of up to 10 000 atoms, with validated paths for larger systems and trajectory data. The environment is designed to be auditable by industrial quality systems and reproducible for academic peer review.

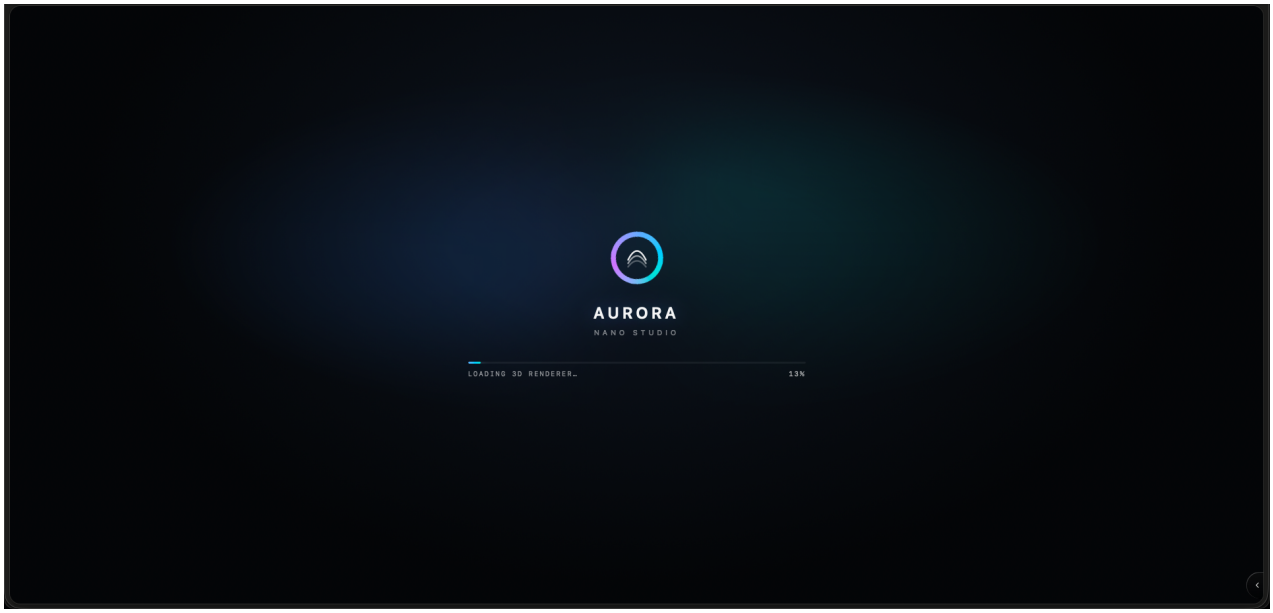
1 Introduction

Atomistic simulation has become central to research in condensed-matter physics, computational chemistry, materials design, pharmaceuticals, semiconductor engineering, and catalysis [12, 13]. First-principles density-functional calculations, classical molecular dynamics, Monte Carlo sampling, and coarse-grained methods together generate structural, dynamical, and electronic data at scales that range from tens of atoms in a quantum-chemical benchmark to tens of millions of atoms in a production molecular dynamics run. The analysis of this data typically requires a combination of crystallographic software, diffraction codes, density-functional post-processing tools, trajectory analysis libraries, and visualization environments—each with its own file conventions, coordinate conventions, and periodicity assumptions.

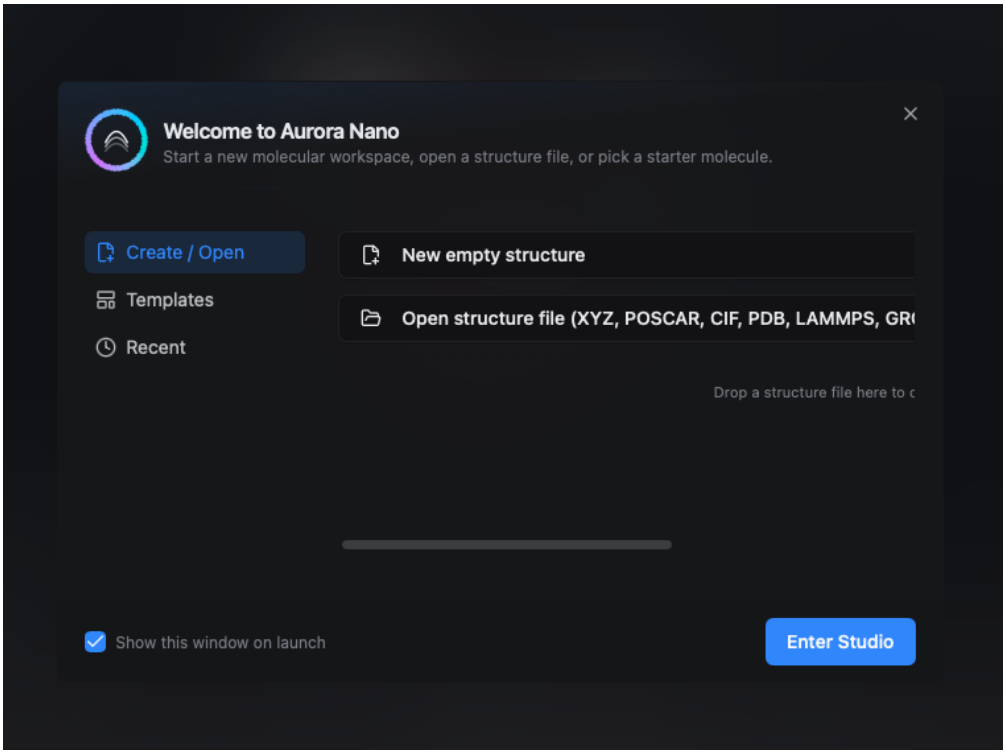
Aurora Nano Studio consolidates this analysis surface into a single environment without replacing or re-implementing the simulation engines themselves. The design principle is analysis proximity: every operation acts on the output of an existing solver rather than substituting for it. The environment reads structure files in crystallographic, molecular, and simulation formats; builds modifier pipelines that process structural data deterministically; supports volumetric scalar fields from density-functional calculations; prepares surfaces and interfaces for subsequent calculation; generates powder-diffraction patterns and refinement diagnostics; visualizes electronic structure outputs; and prepares simulation input decks and molecular systems for downstream engines.

This white paper describes the scientific methods implemented in Aurora Nano Studio. Section organisation follows the natural workflow: crystallographic foundations, chemical bonding, distribution functions and static correlations, dynamical analysis, defect and microstructure analysis, volumetric fields, surface and interface science, phase stability, powder diffraction, electronic structure, and molecular simulation preparation. Section 22 summarises validation evidence and known limitations. Section 23 maps the methods to application domains.

2 Structure Representation and Crystallographic Foundations



(a) Loading screen while the 3D renderer initializes.



(b) Start window for creating or opening a molecular structure.

2.1 Lattice Vectors and the Cell Matrix

A periodic crystalline structure is described by three lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ and a basis of atomic positions. Aurora Nano adopts the right-handed convention used in the International Tables for Crystallography [1]: $\mathbf{a}_1$ is aligned along $+x$, $\mathbf{a}_2$ lies in the xy plane, and $\mathbf{a}_3$ completes the right-handed frame.

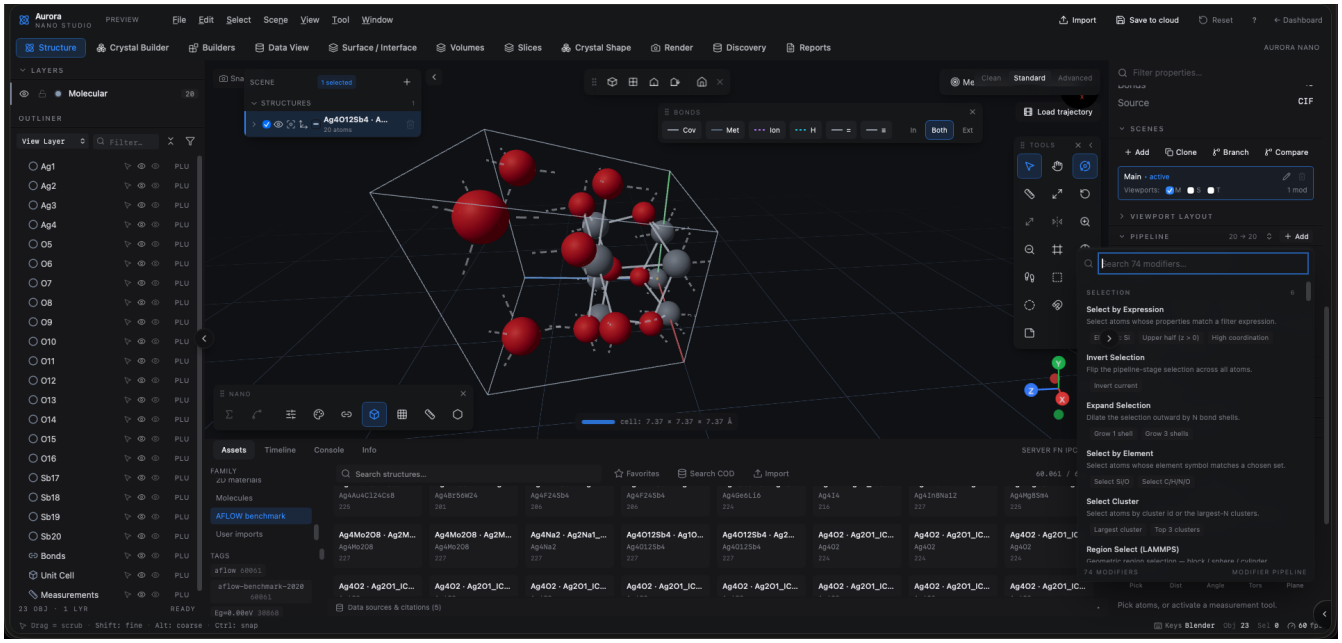


Figure 2: Aurora Nano Studio main interface displaying a molecular structure inside the 3D viewport, with layer controls, structure assets, editing tools, and pipeline modifiers.

Given cell parameters $(a, b, c, \alpha, \beta, \gamma)$ the lattice vectors are:

$$\mathbf{a}_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{a}_2 = \begin{pmatrix} b\cos\gamma \\ b\sin\gamma \\ 0 \end{pmatrix}, \tag{1}$$

$$c_x = c\cos\beta, \qquad c_y = \frac{c(\cos\alpha - \cos\beta\cos\gamma)}{\sin\gamma}, \qquad c_z = \sqrt{\max\bigl(0,\, c^2 - c_x^2 - c_y^2\bigr)}, \tag{2}$$

$$\mathbf{a}_3 = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix}. \tag{3}$$

The cell matrix M stores lattice vectors as columns:

$$M = \begin{bmatrix} a_{1x} & a_{2x} & a_{3x} \\ a_{1y} & a_{2y} & a_{3y} \\ a_{1z} & a_{2z} & a_{3z} \end{bmatrix}. \tag{4}$$

Cartesian and fractional coordinates are related by $\mathbf{r} = M\mathbf{f}$ and $\mathbf{f} = M^{-1}\mathbf{r}$. Fractional wrapping into the unit cell is performed as:

$$f_i' = \bigl((f_i \bmod 1) + 1\bigr) \bmod 1, \tag{5}$$

which is numerically stable for both positive and negative fractional coordinates. The same cell matrix is used for all periodic-boundary operations throughout the analysis pipeline.

2.2 Symmetry, Space Groups, and Wyckoff Positions

Crystallographic symmetry is stored as part of the structure metadata: International number, Hermann–Mauguin symbol, Hall symbol, Hall number, setting, and centering type. The space-group database used in Aurora Nano follows the conventions of [1]. For crystal construction, the environment accepts a space-group number or symbol and a seed basis, then expands the seed atomic positions into complete Wyckoff orbits using the appropriate site-symmetry operations.

Supercell expansion multiplies the lattice vectors by integer factors (n_1, n_2, n_3) and tiles the basis accordingly. Fractional coordinates in the expanded cell are shifted by the appropriate lattice offsets before conversion to Cartesian form.

2.3 Niggli Reduction

For comparison, canonical cell determination, and symmetry analysis, Aurora Nano implements Niggli reduction using the Krivy–Gruber algorithm [2]. The algorithm reduces any lattice to a canonical form in which the three lattice parameters satisfy a set of scalar-product inequalities. The reduced Niggli cell is unique for a given lattice up to orientation, enabling comparison of structures imported from different computational codes. The implementation follows the corrected algorithm discussed in [3].

2.4 Periodic Boundary Conditions

All distance calculations that require periodic boundary conditions use explicit lattice image shifts. For atoms i and j separated by image vector $\mathbf{n} = (n_1, n_2, n_3)$:

$$\Delta\mathbf{r}_{ij}(\mathbf{n}) = \mathbf{r}_i - \Big(\mathbf{r}_j + n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3\Big), \tag{6}$$

$$d_{ij}^2(\mathbf{n}) = \Delta\mathbf{r}_{ij}(\mathbf{n}) \cdot \Delta\mathbf{r}_{ij}(\mathbf{n}). \tag{7}$$

The default image search range is $n_k \in \{-1, 0, +1\}$ along each periodic axis, which is sufficient for bond search at typical covalent-radii cutoffs. For the radial distribution function with orthogonal cells, the minimum-image convention is applied:

$$\Delta x' = \Delta x - a \cdot \mathrm{round}\left(\frac{\Delta x}{a}\right), \tag{8}$$

and analogously for y and z.

3 Chemical Bonding and Molecular Topology

3.1 PBC-Aware Bond Inference

Bond inference in Aurora Nano follows a multi-pass algorithm. In the first pass, all atom pairs (i, j) and image shifts $\mathbf{n}$ are evaluated against the condition:

$$d_{\min}^2 \le d_{ij}^2(\mathbf{n}) \le \big[t\,(r_A + r_B)\big]^2, \tag{9}$$

where r_A and r_B are the covalent radii of the two species from the revised tabulation of [4] and t is a tolerance factor (default $t = 1.15$). Covalent radii cover elements up to atomic number 118.

In the second pass, each candidate bond is classified by bond kind. The classification logic uses element categories (main group, transition metal, rare earth, noble gas) and the Pauling electronegativity difference $|\chi_A - \chi_B|$:

- *Metallic*: both atoms are metals and the electronegativity difference is small.
- *Ionic*: the electronegativity difference exceeds a threshold (typically 1.7).
- *Covalent*: neither metallic nor ionic; the default for main-group bonding.
- *Hydrogen bond*: detected in a separate pass described below.

In the third pass, bond order is assigned for common organic and main-group element pairs using geometric distance bands derived from the bond-length tables of [5]. Bond order 1, 2, or 3 is assigned only when the bond type and distance fall within a calibrated range; ambiguous cases retain bond order 1 to avoid false precision.

3.2 Hydrogen Bond Detection

Hydrogen bonds are detected in a separate pass over donor–hydrogen–acceptor triplets. The donor D and acceptor A must both belong to electronegative elements (N, O, F, S, Cl). The detection condition is:

$$d_{\min} \le \|\mathbf{r}_A - \mathbf{r}_H\| \le d_{\max}, \tag{10}$$

combined with an angular criterion on the D–$H \cdots A$ angle ($\theta \ge 120^\circ$ by default). The hydrogen bond is appended to the bond list with kind `hydrogen` and does not replace the covalent D–H bond already present.

3.3 Effective Coordination Number

Following the formalism of Hoppe [8], Aurora Nano computes the effective coordination number (ECoN) for each atom from the set of bond lengths $\{l_i\}$ to its neighbours:

$$l_{\text{av}} = \frac{\sum_i l_i \, \exp\!\big(1 - (l_i/l_{\min})^6\big)}{\sum_i \exp\!\big(1 - (l_i/l_{\min})^6\big)}, \tag{11}$$

$$\text{ECoN} = \sum_i \exp\!\big(1 - (l_i/l_{\text{av}})^6\big), \tag{12}$$

where $l_{\min}$ is the shortest bond length to any neighbour. ECoN accounts for distorted coordination environments more accurately than a simple integer count and is widely used in inorganic crystal chemistry [8].

3.4 Bond-Valence Sums

Bond-valence analysis [6] provides a local electrostatic check on the chemical reasonableness of a crystal structure. The bond valence for an atom pair is:

$$s_{ij} = \exp\!\left(\frac{R_0 - d_{ij}}{b}\right), \tag{13}$$

where R_0 and $b = 0.37\,\text{Å}$ are tabulated parameters for each pair of element species [6], and d_{ij} is the observed bond length. The bond-valence sum $V_i = \sum_j s_{ij}$ should approach the formal oxidation state of atom i in a well-ordered structure. Significant deviations indicate bond-length strain, disorder, or incorrect symmetry assignment.

3.5 Madelung Electrostatics

The electrostatic energy of an ionic crystal is evaluated using the Ewald summation [7], which decomposes the conditionally convergent lattice sum into a short-range real-space contribution and a rapidly convergent reciprocal-space contribution. For a unit cell with partial charges $\{q_i\}$, the Madelung energy per formula unit is reported in the Reports workspace. This analysis is useful for comparing structural models and for diagnosing ionic polymorph stability.

3.6 SMILES Ingestion and Molecular Graph

For molecular inputs, Aurora Nano accepts SMILES strings [9] with optional implicit hydrogen expansion. The environment uses an RDKit-backed path [10] for canonical SMILES generation, molecular descriptor computation (exact mass, molecular weight), and 2D structure preview. A fallback local parser covers basic organic and inorganic SMILES when the RDKit path is unavailable. The resulting molecular graph is represented as a `Structure` with atoms, Cartesian positions, and an explicit bond list—the same data model used for all crystalline and surface structures throughout the pipeline.

3.7 Polymer Chain Construction

Linear polymer chains are constructed from a repeat-unit SMILES and a degree of polymerisation N_{DP}. The head and tail capping groups can be specified independently. Approximate Cartesian coordinates are generated by sequential placement along a helical backbone, with optional three-dimensional conformer generation delegated to RDKit through the embedded molecule algorithm [10, 11]. The output is a fully bonded `Structure` with per-atom residue tags that preserve the repeat-unit provenance for downstream force-field assignment.

4 Radial and Angular Distribution Functions

4.1 Radial Distribution Function

The pair radial distribution function $g(r)$ is the fundamental structural quantity of a liquid or amorphous solid [12, 14]. It measures the probability density of finding a particle of species B at distance r from a particle of species A, normalised by the mean bulk density. Aurora Nano computes:

$$\rho_B = \frac{N_B}{V}, \tag{14}$$

$$V_{\text{shell}}(r_k) = \frac{4}{3}\pi\left[(r_k + \Delta r)^3 - r_k^3\right], \tag{15}$$

$$g(r_k) = \frac{n(r_k)}{N_A \, V_{\text{shell}}(r_k) \, \rho_B}, \tag{16}$$

where $n(r_k)$ is the observed count of B atoms in shell $[r_k, r_k + \Delta r)$ around all N_A atoms of species A, and V is the cell volume. For same-species pairs the loop avoids double counting and the histogram is corrected accordingly. The minimum-image convention (Eq. 8) is applied for orthogonal cells; full lattice-vector image shifts (Eq. 6) are used for bond and coordination calculations.

4.2 Angular Distribution Function

The angular distribution function measures the probability density of triplet angles θ_{ijk} where j is the central atom and i, k are neighbours within a specified cutoff. Angular distributions are normalised to unit area and are useful for characterising network glasses, hydrogen-bond networks, and coordination polyhedra.

4.3 Bond-Length Distribution

The bond-length distribution records the histogram of all covalent bond lengths in the structure, optionally filtered by element pair. It complements the RDF by focusing on first-shell distances already classified as bonds rather than integrating over all pair distances.

4.4 Static Structure Factor

The static structure factor $S(q)$ is computed from atom positions as a powder-averaged quantity:

$$S(q) = \frac{1}{N} \left| \sum_{j=1}^{N} \exp(i \mathbf{q} \cdot \mathbf{r}_j) \right|^2 , \tag{17}$$

averaged over all wavevector directions at magnitude $q = |\mathbf{q}|$. The structure factor is directly comparable to X-ray and neutron total-scattering measurements and provides the Fourier dual of $g(r)$. For crystalline systems the Debye formula [15] is the foundation of powder diffraction pattern simulation; Section 10 describes the full diffraction workflow.

5 Dynamical Analysis and Trajectory Processing

5.1 Mean-Square Displacement and Diffusion

The mean-square displacement (MSD) characterises atomic mobility in a trajectory of N_{frames} frames:

$$\langle r^2(t) \rangle = \frac{1}{N} \sum_{i=1}^{N} \left| \mathbf{r}_i(t) - \mathbf{r}_i(0) \right|^2 . \tag{18}$$

In the diffusive regime the Einstein relation [16] relates the long-time slope to the self-diffusion coefficient:

$$D = \lim_{t\to\infty} \frac{\langle r^2(t)\rangle}{6t}. \tag{19}$$

Aurora Nano reports the MSD time series for each species separately and includes a linear fit to the diffusive plateau, from which D is extracted in $\mathrm{cm^2\,s^{-1}}$.

5.2 Velocity Autocorrelation Function and Vibrational Density of States

The normalised velocity autocorrelation function (VACF) is:

$$C_{vv}(t) = \frac{\langle \mathbf{v}(0)\cdot\mathbf{v}(t)\rangle}{\langle \mathbf{v}(0)\cdot\mathbf{v}(0)\rangle}, \tag{20}$$

where the average runs over all atoms of a given species and over all time origins in the trajectory. The VACF carries the power spectrum of atomic motion: its Fourier transform gives the vibrational density of states (VDOS) [18]:

$$Z(\nu) = 2\int_0^\infty C_{vv}(t)\, e^{-2\pi i\nu t}\,\mathrm{d}t. \tag{21}$$

Peaks in $Z(\nu)$ correspond to phonon branches and molecular vibrational modes. The VDOS is used to compute the vibrational heat capacity in the harmonic approximation and is compared to infrared and Raman spectra in vibrational analysis workflows.

5.3 Lindemann Melting Index

The Lindemann criterion for melting is expressed through the mean relative displacement of neighbouring atoms [17]:

$$L = \frac{1}{N(N-1)} \sum_i \sum_{j\neq i} \frac{\sqrt{\langle u_{ij}^2\rangle}}{d_{ij}}, \tag{22}$$

where $u_{ij} = |\mathbf{r}_i - \mathbf{r}_j| - \langle |\mathbf{r}_i - \mathbf{r}_j| \rangle$ is the displacement of the interatomic distance from its mean, and d_{ij} is the equilibrium separation. Melting in most materials occurs when $L \approx 0.10$–0.15. Aurora Nano computes L per atom and per species, enabling spatially resolved detection of pre-melting disorder.

5.4 Trajectory Import, Indexing, and Time Averaging

Large trajectory files (LAMMPS dump, XDATCAR, DL_POLY HISTORY, OUTCAR, *vasprun.xml*, extXYZ) are handled through a lazy frame-index scheme. On first access the system constructs an index of byte offsets, frame counts, and atomic-species metadata without loading all frame coordinates. Individual frames are read on demand. This allows interactive analysis of trajectories with 10^5 or more frames without retaining all positions in memory simultaneously. Time averaging over a configurable window of frames is supported for structural analysis modifiers that benefit from reduced thermal noise.

6 Defect Analysis and Local Structure Classification

6.1 Centrosymmetry Parameter

The centrosymmetry parameter (CSP) [19] quantifies the degree of inversion symmetry around each atom and is used to identify defects in otherwise centrosymmetric crystal structures. For an atom with n nearest neighbours, Aurora Nano finds the $n/2$ disjoint pairs (i_k, j_k) that minimise the sum of vector norms:

$$\mathrm{CSP} = \sum_{k=1}^{n/2} \|\mathbf{r}_{i_k} + \mathbf{r}_{j_k}\|^2, \tag{23}$$

where $\mathbf{r}_{i_k}$ and $\mathbf{r}_{j_k}$ are the position vectors of atom i_k and j_k relative to the central atom. Pairs are selected greedily in order of increasing $\|\mathbf{r}_{i_k} + \mathbf{r}_{j_k}\|^2$. CSP is zero for perfectly centrosymmetric environments and increases with displacement from the ideal lattice, making it sensitive to stacking faults, dislocations, grain boundaries, and free surfaces.

6.2 Wigner–Seitz Vacancy and Interstitial Mapping

Wigner–Seitz analysis identifies vacancies and interstitials by comparing a defective snapshot against a reference perfect-crystal configuration. Each atom in the reference is assigned a Wigner–Seitz cell; current-frame atoms falling in a cell with no reference atom are classified as interstitials, and reference cells containing no current-frame atom are classified as vacancies. The analysis uses a spatial grid to avoid an $O(N^2)$ search.

6.3 Atomic Strain and Non-Affine Displacement

The local deformation gradient tensor at atom i is obtained by comparing reference-frame neighbor vectors $\{\Delta\mathbf{r}_j^{\mathrm{ref}}\}$ to current-frame neighbor vectors $\{\Delta\mathbf{r}_j^{\mathrm{cur}}\}$ [20]:

$$X = \sum_j \Delta\mathbf{r}_j^{\mathrm{ref}} \otimes \Delta\mathbf{r}_j^{\mathrm{ref}}, \qquad Y = \sum_j \Delta\mathbf{r}_j^{\mathrm{cur}} \otimes \Delta\mathbf{r}_j^{\mathrm{ref}}, \qquad F = YX^{-1}. \tag{24}$$

The Lagrangian strain tensor and volumetric strain follow:

$$\eta = \tfrac{1}{2}(F^T F - I), \qquad \eta_{\mathrm{vol}} = \det(F) - 1. \tag{25}$$

The non-affine displacement D^2_{min} measures how much the local deformation deviates from a pure affine mapping:

$$D^2_{\mathrm{min}} = \sum_j \|\Delta\mathbf{r}_j^{\mathrm{cur}} - F\Delta\mathbf{r}_j^{\mathrm{ref}}\|^2. \tag{26}$$

D^2_{min} is large at shear-transformation zones in metallic glasses [20] and at dislocation cores in crystals. Atoms with fewer than four non-coplanar neighbours receive `NaN` to preserve data integrity.

6.4 Common-Neighbor Analysis

Common-neighbor analysis (CNA) [21] classifies atoms by the connectivity of their shared neighbours. For each bonded pair (i, j), the algorithm records the number of common neighbours, the number of bonds among those neighbours, and the length of the longest chain of such bonds—giving the triple (n_c, n_b, n_l). Characteristic fingerprints distinguish face-centred cubic (fcc), hexagonal close-packed (hcp), body-centred cubic (bcc), and icosahedral environments. The Ackland–Jones variant [22] uses angular histograms over all neighbour pairs and is provided as an alternative for high-temperature or defective structures.

6.5 Polyhedral Template Matching

Polyhedral template matching (PTM) [23] provides robust crystal-structure identification that is less sensitive to thermal vibrations than CNA. The method matches the local coordination polyhedron of each atom against template polyhedra for fcc, hcp, bcc, simple cubic, diamond cubic, and graphene structures. The matching is performed using the RMSD between the current local environment and the ideal template after optimal rotation, returning a structure type and a numerical RMSD that serves as a confidence measure.

6.6 Steinhardt Bond-Orientational Order Parameters

The Steinhardt bond-orientational order parameters q_l [24] project the local bond network onto spherical harmonics of degree l:

$$Q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_{lm}(\hat{r}_{ij}) \,, \tag{27}$$

$$q_l(i) = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |Q_{lm}(i)|^2 \right)^{1/2} , \tag{28}$$

where $N_b(i)$ is the number of neighbours within the cutoff and Y_{lm} are the real spherical harmonics. q_4 and q_6 are diagnostic of fcc, hcp, bcc, icosahedral, and liquid environments and are widely used to detect crystal nucleation in molecular dynamics simulations of crystallisation.

6.7 Additional Classification Methods

Aurora Nano also implements:

- *CHILL+ ice classification* [25]: identifies cubic ice, hexagonal ice, interfacial ice, and liquid water from local structure.
- *Diamond structure identification*: assigns cubic or hexagonal diamond character based on second-shell coordination geometry.
- *Slip vector analysis* [26]: computes the Burgers-vector-like displacement of each atom relative to its neighbours in a reference configuration, highlighting dislocation cores and stacking faults.
- *Planar fault analysis*: detects intrinsic and extrinsic stacking faults from layer-sequence mismatches.

- *Dislocation extraction (MVP)*: extracts approximate dislocation lines from Burgers-circuit analysis on a per-atom basis.
- *Grain segmentation (MVP)*: clusters atoms by orientational order into polycrystalline grains.
- *Voronoi topology (MVP)*: classifies coordination polyhedra by Voronoi face count following Voronoi [27] and subsequent computational implementations [28].

Items marked MVP are available in the modifier registry as initial implementations with stated limitations on robustness for all system types.

6.8 Cluster Analysis

Aurora Nano's cluster modifier uses a disjoint-set union (union-find) data structure with path compression and union-by-rank [29]. Connectivity is derived from the bond graph or from a cutoff-based spatial grid. Connected components are ordered by size, and components below a minimum-size threshold are assigned to a background class. The result—a per-atom cluster ID and optional cluster-averaged color—feeds downstream selections such as "isolate the largest cluster" or "remove clusters smaller than N atoms".

7 Volumetric Field Analysis and Electron Density

7.1 Supported Scalar Field Types

Aurora Nano loads and analyses periodic scalar fields on a rectangular grid aligned with the unit cell. Supported field types include:

- Electron charge density $\rho(\mathbf{r})$ (CHGCAR, CUBE, XSF, PARCHG, OpenDX, VTI).
- Electrostatic potential $V(\mathbf{r})$ (LOCPOT, OpenDX).
- Partial charge density (PARCHG, CUBE).
- Spin density $\rho_{\uparrow}(\mathbf{r}) - \rho_{\downarrow}(\mathbf{r})$ (second component of CHGCAR when spin-polarised).
- Charge difference density $\Delta\rho = \rho_{\text{system}} - \sum_i \rho_{\text{atom}\,i}$.
- Electron localization function $\chi(\mathbf{r})$ (described below).
- Local density of states (L/PDOS energy-resolved grids where available).

7.2 Derived Fields

Aurora Nano can compute derived scalar fields from a loaded primary field:

- Gradient magnitude: $|\nabla\rho|$, useful for identifying bonding and bonding regions.
- Laplacian: $\nabla^2\rho$, central to Bader charge analysis [31].
- Kinetic energy density: $\tau(\mathbf{r}) = \frac{1}{2}\sum_i |\nabla\psi_i(\mathbf{r})|^2$ (from orbitals when available).
- Model electron density: constructed from atomic electron density radial functions as a preview when no DFT field is available.

7.3 Electron Localization Function

The electron localization function (ELF) [30] is a dimensionless scalar field in $[0, 1]$ that measures the degree of electron localization relative to a homogeneous electron gas reference:

$$D(\mathbf{r}) = \tau(\mathbf{r}) - \frac{|\nabla\rho(\mathbf{r})|^2}{4\rho(\mathbf{r})}, \tag{29}$$

$$D_h(\mathbf{r}) = \tfrac{3}{5}(6\pi^2)^{2/3}\rho(\mathbf{r})^{5/3}, \tag{30}$$

$$\chi(\mathbf{r}) = \left[1 + \left(\frac{D(\mathbf{r})}{D_h(\mathbf{r})}\right)^2\right]^{-1}. \tag{31}$$

Values near 1.0 indicate covalent bonding or lone pairs; values near 0.5 indicate metallic or delocalised electron behaviour. Aurora Nano renders the ELF as an isosurface using the domain-specific presets: bonding character ($\chi = 0.5$), shell structure ($\chi = 0.75$), and lone-pair basins ($\chi = 0.85$).

7.4 Tricubic Interpolation and Upsampling

Interactive iso-surface rendering and line-profile extraction require smooth interpolation of the discrete grid. Aurora Nano uses tricubic interpolation on the $3 \times 3 \times 3$ neighbourhood of each query point, providing C^1 continuity across the field without the Gibbs-like artefacts of trilinear methods. An upsampling option allows the grid to be resampled at double the original resolution before rendering, improving surface quality for coarsely sampled DFT outputs.

7.5 Marching-Cubes Iso-Surface Extraction

Iso-surfaces are extracted using the marching-cubes algorithm of Lorensen and Cline [33], which decomposes the grid into $2 \times 2 \times 2$ voxels and triangulates the surface within each voxel by table lookup. The resulting triangle mesh is normalised and rendered in the three-dimensional viewport. Users can specify one or more iso-levels with independent colors and opacity values, enabling stacked positive/negative density views and shell-by-shell ELF inspection. Interactive memory budgets constrain the number of triangles generated to avoid browser memory exhaustion.

7.6 Two-Dimensional Slice Engine

The slice engine extracts a two-dimensional map by sampling the scalar field on a plane defined by one of four methods:

1. *Axis-aligned*: a plane perpendicular to a Cartesian axis at a specified fractional coordinate.
2. *Miller plane*: a plane perpendicular to the reciprocal-lattice vector $\mathbf{G}_{hkl} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3$.
3. *Vector plane*: a plane defined by two arbitrary vectors and an origin in fractional coordinates.
4. *Projection*: integration of the field along a chosen axis, yielding a projected density map.

The extracted 2D map is rendered as a false-colour image with a user-selected colormap. Overlaid marching-squares contours mark constant-density lines. A line profile tool samples the map along an arbitrary direction. Peak search identifies local maxima in 1D or 2D. SVG and CSV exports are provided for publication-quality reproduction.

8 Surface and Interface Science

8.1 Slab Construction from Miller Indices

Surfaces are prepared by cleaving a bulk structure along a specified Miller plane (hkl). The algorithm: (1) determines the surface-plane unit vectors $\mathbf{u}$ and $\mathbf{v}$ lying in the (hkl) plane, (2) constructs a minimal orthorhombic or hexagonal surface supercell by searching for the smallest repeat vectors, (3) tiles the bulk cell to fill the slab dimensions, (4) wraps atoms into the slab, (5) applies a vacuum layer of specified thickness along the surface normal, and (6) optionally passivates dangling bonds by adding hydrogen or other specified species at a fixed bond length and angle. The resulting structure is stored with surface-normal metadata and is compatible with VASP and Quantum ESPRESSO input generation.

8.2 Lattice Matching for Interfaces

Interface construction requires finding supercell pairs of two surface unit cells that are commensurate. Following the method of Zur and McGill [34], Aurora Nano searches over integer supercell matrices $(m_{11}, m_{12}; m_{21}, m_{22})$ for each surface and identifies pairs with small in-plane lattice mismatch. The mismatch strain is:

$$\varepsilon = \frac{a_{\mathrm{film}} - a_{\mathrm{sub}}}{a_{\mathrm{sub}}}, \tag{32}$$

and the search minimises $|\varepsilon|$ subject to a user-specified tolerance and a maximum supercell area. The resulting matched supercell is used as the lateral repeat for the interface slab. For epitaxial systems the elastic mismatch energy per unit area is estimated as:

$$E_{\mathrm{el}} = \frac{Y\, h\, \varepsilon^2}{1 - \nu}, \tag{33}$$

where Y is the biaxial modulus, h is the film thickness, and ν is the Poisson ratio of the film material.

8.3 Interface Adhesion Energy

The adhesion energy per unit area of an interface between two slabs is:

$$E_{\mathrm{adh}} = \frac{E_{\mathrm{interface}} - E_{\mathrm{slab\,1}} - E_{\mathrm{slab\,2}}}{A}, \tag{34}$$

where $E_{\mathrm{interface}}$ is the total energy of the combined system, $E_{\mathrm{slab\,1}}$ and $E_{\mathrm{slab\,2}}$ are the energies of the isolated slabs in the same supercell, and A is the in-plane area. Aurora Nano computes E_{adh} from energies attached to the three structures and generates a separation-energy curve as a function of the gap between the two surfaces, from which the interface binding well is characterised.

8.4 Registry Scans

For van der Waals heterostructures and weakly bonded interfaces, the energy landscape as a function of lateral shift (registry) is important. Aurora Nano generates a grid of rigid-shift configurations, attaches calculator-derived energies from a scan table, and renders the registry energy map in the two-dimensional lateral plane.

8.5 Adsorption Site Analysis

Adsorption site identification searches the surface for hollow, bridge, and top sites using a Voronoi tessellation of the surface-layer atoms projected onto the surface plane [27]. Each identified site is ranked by a Lennard-Jones-style preview energy [35]:

$$E_{\mathrm{LJ}}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right], \tag{35}$$

summed over surface atoms, providing a rapid ordering of site candidates before expensive density-functional calculations. Calculator-derived adsorption energies can be imported from a CSV or JSON table and displayed alongside the preview ranks.

8.6 Computational Input Generation

Aurora Nano generates VASP [36] and Quantum ESPRESSO [37] input decks from the active structure, including POSCAR/KPOINTS/POTCAR manifest, `pw.x` input, and pseudopotential manifest. The generated decks include lattice parameters, atomic positions, k-point sampling, and commonly used DFT convergence parameters as defaults that the user can customise. Adsorption scan bundles and interface separation scan bundles are exported as structured JSON archives containing all structures, metadata, and submission scripts.

9 Phase Stability and Thermodynamics

9.1 Convex Hull Construction

The thermodynamic stability of a compound is assessed by its position on the convex hull of formation energies as a function of composition [41]. For a binary system A–B, the formation energy per atom is:

$$\Delta H_f(x_B) = H_{\mathrm{compound}}(x_B) - (1 - x_B)\, H_A - x_B\, H_B, \tag{36}$$

where x_B is the B-atom fraction and H_A, H_B are the reference energies of the elemental solids. A compound is thermodynamically stable (at 0 K) if it lies on the lower convex hull of the ΔH_f vs. x_B plot. The energy above hull ΔH_{hull} quantifies the degree of metastability.

For ternary and higher-order systems, the convex hull is constructed in composition simplex space. Aurora Nano visualises binary hulls as ΔH_f vs. x_B line plots and ternary hulls as colour-mapped triangular plots, with selectable entries, stable-phase tie-lines, and nearest-stable-phase decomposition routes.

9.2 Chemical Potential Diagrams

Chemical potential diagrams map the thermodynamic stability window of a target compound as a function of the reservoir chemical potentials $\Delta\mu_A$, $\Delta\mu_B$ (referenced to the elemental chemical potential, so $\Delta\mu \leq 0$). For each pair of stable phases in a binary system, the coexistence chemical-potential vertex is obtained by solving the 2×2 linear system via Cramer's rule:

$$\det = x_A^{(2)} - x_A^{(1)}, \tag{37}$$

$$\Delta\mu_A = \frac{h^{(1)}\, x_A^{(2)} - h^{(2)}\, x_A^{(1)}}{\det}, \tag{38}$$

$$\Delta\mu_B = \frac{(1 - x_A^{(1)})\, h^{(2)} - (1 - x_A^{(2)})\, h^{(1)}}{\det}, \tag{39}$$

where $x_A^{(k)}$ is the A-atom fraction of phase k and $h^{(k)}$ is its formation energy per atom. The stability polygon in $(\Delta\mu_A, \Delta\mu_B)$ space is the convex region satisfying all stability constraints simultaneously, i.e. $x_A^{(k)}\Delta\mu_A + x_B^{(k)}\Delta\mu_B \leq h^{(k)}$ for all competing phases k. This polygon is computed by finding all pairwise constraint-line intersections and retaining only those vertices satisfying all constraints with $\Delta\mu_A \leq 0$ and $\Delta\mu_B \leq 0$. The polygon is then sorted into convex order.

Phase stability assessments of this type are a key tool in the design of synthesis conditions, battery electrode materials, and corrosion-resistant alloys [42, 43].

10 Powder Diffraction and Crystallographic Reporting

10.1 Powder Diffraction Pattern Simulation

X-ray powder diffraction patterns are simulated from the crystal structure using the kinematic approximation. Bragg reflections occur at positions given by:

$$2\, d_{hkl}\, \sin\theta_{hkl} = n\lambda, \tag{40}$$

where d_{hkl} is the d-spacing of the (hkl) planes, θ_{hkl} is the Bragg angle, and λ is the X-ray wavelength (Cu Kα_1: 1.5406 Å by default). The structure factor for reflection (hkl) is:

$$F_{hkl} = \sum_j f_j(Q)\, e^{2\pi i(hx_j + ky_j + lz_j)}\, T_j(Q), \tag{41}$$

where $f_j(Q)$ is the atomic form factor of species j evaluated at $Q = 4\pi \sin\theta/\lambda$, (x_j, y_j, z_j) are fractional coordinates, and $T_j(Q) = e^{-B_j Q^2/(16\pi^2)}$ is the Debye–Waller temperature factor with isotropic displacement parameter B_j. The intensity $I_{hkl} \propto |F_{hkl}|^2 \cdot m_{hkl} \cdot \mathrm{LP}(\theta)$, where m_{hkl} is the reflection multiplicity and LP is the Lorentz–polarisation factor.

10.2 Peak Shape and Pattern Profile

The calculated pattern is convolved with a pseudo-Voigt peak-shape function [47]:

$$\mathrm{pV}(x) = \eta\, L(x) + (1-\eta)\, G(x), \tag{42}$$

where L is a Lorentzian, G is a Gaussian, and η is the mixing parameter determined from the Caglioti relation [47] as a function of θ.

10.3 Rietveld Refinement and Related Methods

Aurora Nano's powder panel supports comparison of a simulated pattern against an observed (experimental) trace. The scientific foundation includes:

- *Rietveld refinement* [44]: least-squares minimisation over structural and profile parameters.
- *Pawley fitting* [45]: profile-only refinement without a structural model, yielding unit-cell parameters and reflection intensities.
- *Le Bail decomposition* [46]: iterative intensity extraction without a structural model, useful for phase identification from powder data alone.

The environment provides a diagnostic panel showing the difference pattern, R-factors, and per-reflection statistics. Peak matching against observed-trace peaks identifies reflections that exceed a configurable intensity and d-spacing tolerance.

10.4 Crystal Morphology

Crystal morphology is predicted from the Miller face list using the Bravais–Friedel–Donnay–Harker (BFDH) rule [48]: face importance is proportional to the d-spacing of the (hkl) family. Aurora Nano generates the morphology mesh from the specified face set, computes face areas and interfacial angles, expands the set to include symmetry-equivalent faces, and visualises the habit as a closed polyhedral mesh with optional per-face color and opacity. Twin-domain visualization superimposes a twinned orientation related by a specified twin law.

10.5 Patterson Fourier Maps

The Patterson function [49]:

$$P(\mathbf{u}) = \frac{1}{V} \sum_{hkl} |F_{hkl}|^2 \cos(2\pi\, \mathbf{h} \cdot \mathbf{u}) \tag{43}$$

is a peak-finding tool for locating heavy atoms in a crystal structure without phase information. Peaks in $P(\mathbf{u})$ correspond to interatomic vectors; their heights are proportional to $Z_i\, Z_j$ products. Aurora Nano computes the Patterson map from the structure-factor amplitudes of the current structure and renders it on a grid via the slice engine.

11 Electronic Structure Analysis

11.1 Band Structure

The electronic band structure $\varepsilon_n(\mathbf{k})$ is the fundamental output of a density-functional calculation for a periodic system [38, 39]. Aurora Nano parses VASP EIGENVAL files and plots band energies as functions of $\mathbf{k}$-path position. The Fermi energy E_F is used as the zero reference. Band-gap determination identifies the conduction-band minimum (CBM) and valence-band maximum (VBM), classifies the gap as direct or indirect, and reports the gap value in eV.

11.2 Density of States

The density of states (DOS) is parsed from VASP DOSCAR files, which provide the total DOS and optionally the projected (partial) DOS (PDOS) decomposed by atom type, orbital angular momentum, and spin channel. The DOS at the Fermi level $N(E_F)$ is reported as a metallic-character indicator. Fermi-surface crossings are identified from bands that cross E_F. The integrated DOS from the band bottom to E_F gives the total number of occupied states per formula unit.

11.3 Brillouin Zone and k-Path Generation

Standard high-symmetry k-paths for band structure calculations are generated following the conventions of [40], which define reciprocal-space paths through the irreducible Brillouin zone for each Bravais lattice type. Aurora Nano provides standard paths, primary paths (shortest connected path through high-symmetry points), and Γ-centred paths. The k-path is exported as a VASP `KPOINTS` file in line mode or as a Quantum ESPRESSO K_POINTS block.

11.4 Fermi Surface

For metallic systems, a coarse Fermi-surface mesh is constructed by identifying bands within a thermal window $\pm k_B T$ of E_F and sampling the $(\mathbf{k}, \varepsilon)$ surface using a uniform $\mathbf{k}$-grid. The resulting mesh triangle count and surface area serve as a quick metallic-character estimate pending a full Fermi-surface calculation.

12 Molecular Simulation Preparation

12.1 Polymer Builder

The polymer builder accepts a repeat-unit SMILES, a degree of polymerisation N_{DP}, and optional head/tail capping SMILES. The chain is constructed by sequential covalent linkage through the terminal reactive atoms of the repeat unit. Three-dimensional conformers are generated using the distance-geometry plus force-field optimisation pipeline in RDKit [10, 11]. The resulting chain structure carries per-atom residue indices for downstream force-field typing.

12.2 Simulation Box Packing

Molecular systems for liquid-state or solution MD are prepared by packing multiple molecular species into a simulation box at a target density or with specified counts. The packing algorithm places molecules

at pseudo-random positions with minimum-distance constraints to avoid hard-core overlap—following the conceptual framework of Packmol [54]. Server-side execution with Packmol or an equivalent packing engine is supported through the compute-provider interface for production runs.

12.3 Force-Field Assignment

Force-field assignment maps each atom type in the molecular graph to a parameter set from a chosen force field. Aurora Nano supports three assignment paths:

- *OpenFF SMIRNOFF* [52]: SMARTS-based typing from the Open Force Field Toolkit; covers small organic molecules and peptides with modern, reproducible parameters.
- *OPLS-AA/Foyer* [50]: atom-type assignment via SMARTS rules, widely used for biomolecular and small-molecule MD in GROMACS, LAMMPS, and OpenMM.
- *GAFF/AmberTools* [51]: generalised Amber force field for small drug-like molecules; requires an external AmberTools installation for antechamber-based charge computation.

The assignment report lists typed and untyped atoms, partial charges (where available), and any missing parameters, enabling a review step before simulation deck generation.

12.4 Simulation Deck Generation

Aurora Nano generates ready-to-run input decks for OpenMM [53] (XML system file), LAMMPS (data file and input script), and GROMACS (top and gro files where supported). The generated decks are exported as an artifact bundle together with the typed structure, force-field parameters, and a metadata manifest. Simulation parameters (temperature, pressure ensemble, timestep, cutoffs) are set to physically reasonable defaults and reviewed in the panel before export.

12.5 Heavy Trajectory Analysis

For trajectory observables that are too computationally intensive for interactive browser analysis—including hydrogen-bond network lifetimes, Van Hove correlation functions, dielectric susceptibility, ionic conductivity via Green–Kubo integration [18], vibrational IR/Raman spectra from the VACF, and spatial distribution functions—Aurora Nano delegates to server-side MDAnalysis [55] or MDTraj [56] handlers. Results are returned as typed plot and table artifacts that are displayed in the analysis viewport using the same plot panels as browser-side analyses.

13 Professional Platform Boundary and Scientific Disclosure

Aurora Nano Studio is a commercial scientific platform. This white paper therefore separates two categories of information. The first category is the public scientific basis: equations, observable definitions, validation expectations, and interpretation logic that a trained scientist or engineer needs in order to trust the platform. The second category is the proprietary production boundary: internal routing heuristics, ranking functions, cache policies, partner-specific calibration constants, fault-recovery decisions, and implementation details that are not required for scientific interpretation and are not disclosed in this document.

The platform posture is intentionally similar to professional analytical instrumentation. A modern diffractometer, spectrometer, or electron microscope can publish the physical measurement model, calibration traceability, uncertainty language, and output definitions without disclosing every firmware routine, detector correction, or vendor-specific optimization. Aurora Nano follows the same principle. The reader should be able to understand what a reported quantity means, how it relates to accepted chemistry and physics, and what validation boundary supports it, while the product preserves commercial know-how.

13.1 Disclosure Layers

Layer	Exposed in this paper	Retained as proprietary platform know-how
Scientific model	Coordinate systems, metrics, observables, physical definitions, and equation-level interpretation.	Internal numerical tuning, provider selection heuristics, adaptive thresholds, and performance-specific implementation details.
Analysis output	Definition of each scalar, vector, histogram, surface, or report artifact returned to the user.	Internal object layouts, cache invalidation choices beyond the public contract, and code-level execution pathways.
Validation	Test classes, benchmark budgets, error handling principles, and known scientific limitations.	Private test corpora, partner data, synthetic stress suites, and commercial acceptance gates.
Enterprise control	Audit concepts, lineage objects, governed artifact policy, and deployment-neutral compute boundaries.	Deployment-specific security policies, tenant-specific routing rules, billing weights, and operational playbooks.

This separation protects the product while strengthening the technical narrative. It demonstrates that Aurora Nano is not a black-box renderer, but it also avoids turning the white paper into an engineering manual. The public contract is the scientific and operational meaning of the result; the private contract is how the production platform delivers that result under commercial constraints.

13.2 Scientific Contract of a Professional Platform

For a reported observable Q, Aurora Nano treats the professional contract as a tuple

$$\mathcal{C}(Q) = (\mathcal{D},\, \mathcal{M},\, \mathcal{P},\, \mathcal{U},\, \mathcal{L})\,, \tag{44}$$

where $\mathcal{D}$ is the data domain, $\mathcal{M}$ is the mathematical definition, $\mathcal{P}$ is the physical interpretation, $\mathcal{U}$ is the uncertainty or limitation statement, and $\mathcal{L}$ is the lineage record that connects the result to source files, parameters, and compute state. This contract is deliberately broader than a numerical output. It gives industrial teams the context required to decide whether a result is appropriate for screening, design review, publication support, or regulated documentation.

The same principle applies to derived artifacts. A color-coded atomistic structure, for example, is not only a picture. It is a rendering of a typed attribute field a_i over atomic sites $i = 1, \ldots, N$, together with a color map, selection rule, provenance record, and operation definition. A diffraction plot is not only an

image; it is a sampled function $I(2\theta)$ with a wavelength, profile model, reflection list, and normalization convention. A volumetric isosurface is not only a mesh; it is the level set of a scalar field $\rho(\mathbf{r})$ on a periodic grid, transformed through an interpolation and surface extraction model.

13.3 Professional Language for Proprietary Methods

The product should not describe sensitive internals as step-by-step recipes. Instead, the documentation should use professional disclosure language:

- *Definition*: what the observable means mathematically.
- *Scope*: which inputs and physical regimes the observable supports.
- *Assumption*: which approximations are present, such as orthogonal minimum-image RDF or selected bond-order tables.
- *Validation*: what is checked, benchmarked, or guarded.
- *Boundary*: which details are implementation-specific and intentionally not disclosed.

This convention lets Aurora Nano sound like a professional scientific instrument while retaining the proprietary boundary that protects a product. The following sections add deeper mathematical, chemical, and physical derivations in that spirit: enough for expert confidence, not enough to expose private implementation mechanics.

14 Extended Mathematical Derivations for Periodic Geometry

14.1 Metric Tensor and Cell Volume

The cell matrix M introduced in Section 2 defines a metric tensor

$$G = M^T M, \tag{45}$$

so that the squared length of a displacement written in fractional coordinates $\Delta\mathbf{f}$ is

$$\|\Delta\mathbf{r}\|^2 = \Delta\mathbf{f}^T G\, \Delta\mathbf{f}. \tag{46}$$

Expanding Eq. 45 in crystallographic parameters gives

$$G = \begin{bmatrix} a^2 & ab\cos\gamma & ac\cos\beta \\ ab\cos\gamma & b^2 & bc\cos\alpha \\ ac\cos\beta & bc\cos\alpha & c^2 \end{bmatrix}. \tag{47}$$

The cell volume is

$$V = |\det M| = abc\sqrt{1 + 2\cos\alpha\cos\beta\cos\gamma - \cos^2\alpha - \cos^2\beta - \cos^2\gamma}. \tag{48}$$

This quantity is used throughout density, RDF, reciprocal-space, and phase-stability calculations. The volume formula also exposes a useful validation condition: a physically valid three-dimensional unit cell

requires $V > 0$ and a positive-definite metric tensor. Near-zero V is not merely a numerical inconvenience; it indicates a degenerate coordinate system that cannot support reliable periodic analysis.

14.2 Reciprocal Lattice and Miller Planes

The reciprocal lattice matrix B is defined by

$$B = 2\pi M^{-T}, \tag{49}$$

with columns $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$ satisfying $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$. A Miller index triplet (hkl) defines a reciprocal vector

$$\mathbf{g}_{hkl} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3. \tag{50}$$

The interplanar spacing is therefore

$$d_{hkl} = \frac{2\pi}{\|\mathbf{g}_{hkl}\|}. \tag{51}$$

Equations 49–51 connect the real-space structure to diffraction, morphology, Brillouin-zone construction, and reciprocal-space visualization. They also provide a compact way to validate imported cells: a reported d-spacing table must be consistent with the same M that drives periodic distances.

14.3 Fractional Minimum Images in General Cells

For a displacement $\Delta\mathbf{f} = \mathbf{f}_i - \mathbf{f}_j$ in fractional coordinates, the exact periodic image distance for a general cell is the integer minimization

$$d_{ij}^2 = \min_{\mathbf{n}\in\mathbb{Z}^3} (\Delta\mathbf{f} + \mathbf{n})^T G (\Delta\mathbf{f} + \mathbf{n}). \tag{52}$$

Equation 52 is the mathematical target behind all lattice-aware neighbor definitions. In an orthogonal cell the metric tensor is diagonal and the familiar component-wise rounding rule follows immediately. In a skewed cell, however, the closest image can depend on off-diagonal metric terms such as $ab\cos\gamma$. This is why Aurora Nano distinguishes explicitly between orthogonal minimum-image RDF and full lattice-vector image handling in bond and coordination paths.

The public scientific interpretation is simple: a periodic neighbor is the image that minimizes the metric length, not necessarily the image obtained by independently wrapping each coordinate. The production implementation can accelerate that search in proprietary ways, but the scientific contract remains Eq. 52.

14.4 Tensor Transformation Between Fractional and Cartesian Frames

Some atomistic quantities are scalar and invariant under coordinate transformation. Others, including strain, diffusion, and covariance tensors, must be transformed consistently. For a second-rank tensor A_f represented in fractional basis coordinates, the corresponding Cartesian tensor is

$$A_r = M A_f M^T. \tag{53}$$

For a reciprocal-space tensor K_f, the corresponding Cartesian form uses the reciprocal basis:

$$K_r = M^{-T} K_f M^{-1}. \tag{54}$$

These transformations are important when comparing anisotropic diffusion, elastic strain, thermal displacement ellipsoids, reciprocal grids, and Brillouin-zone paths. A visually correct structure can still produce a physically incorrect tensor if the basis transformation is not handled consistently.

14.5 Cell Deformation and Finite Strain

When a trajectory or relaxation changes the cell from M_0 to M_t, the cell-level deformation gradient is

$$F_{\text{cell}} = M_t M_0^{-1}. \tag{55}$$

The corresponding Green-Lagrange strain is

$$E_{\text{cell}} = \frac{1}{2}\left(F_{\text{cell}}^T F_{\text{cell}} - I\right). \tag{56}$$

For small deformations $F_{\text{cell}} \approx I + \nabla \mathbf{u}$, Eq. 56 reduces to the familiar infinitesimal strain tensor

$$\varepsilon = \frac{1}{2}\left(\nabla \mathbf{u} + \nabla \mathbf{u}^T\right). \tag{57}$$

Aurora Nano can therefore present deformation consistently across scales: local non-affine atomic strain from neighbor fits, global cell strain from lattice vectors, and visual inspection from atomic displacement fields.

15 Extended Chemistry and Bonding Derivations

15.1 Bond-Length Residuals and Chemical Plausibility

The bond-inference section describes how bonds are classified. A complementary chemistry diagnostic is the dimensionless bond-length residual

$$\epsilon_{ij}^{\text{bond}} = \frac{d_{ij} - d_{AB}^0}{d_{AB}^0}, \tag{58}$$

where d_{AB}^0 is a reference distance for species pair A–B, typically derived from covalent radii or tabulated crystallographic bond lengths [4, 5]. A residual near zero indicates a distance consistent with the reference environment. A positive residual indicates stretching or possible weak interaction. A negative residual indicates compression, high bond order, unusual oxidation state, or a geometry that requires expert review.

A structure-level residual can be summarized as

$$R_{\text{bond}} = \sqrt{\frac{1}{N_b} \sum_{(i,j)\in\mathcal{B}} \left(\epsilon_{ij}^{\text{bond}}\right)^2}, \tag{59}$$

where $\mathcal{B}$ is the accepted bond set and N_b is the number of bonds. This is not a replacement for chemical judgment. It is an inspection aid: high residuals direct the user toward bonds, coordination polyhedra, or imported coordinates that deserve closer review.

15.2 Bond-Valence Residual and Oxidation-State Consistency

Equation 13 defines the individual bond valence s_{ij}. For atom i, the deviation from a target formal valence z_i is

$$\Delta V_i = \sum_j s_{ij} - z_i. \tag{60}$$

The crystal-wide bond-valence mismatch can be written as

$$\chi^2_{\mathrm{BVS}} = \frac{1}{N} \sum_{i=1}^{N} (\Delta V_i)^2 . \tag{61}$$

This expression has practical value in oxide, halide, sulfide, and battery materials. If a nominal oxidation-state assignment produces large χ^2_{BVS}, the issue may be a distorted coordination environment, mixed valence, disorder, vacancy concentration, incorrect symmetry, or an imported model that does not correspond to a chemically reasonable minimum. Aurora Nano presents such metrics as diagnostics, not as automatic truth.

15.3 Ionic, Covalent, Metallic, and Hydrogen-Bonding Indicators

Chemical bonding is not a discrete switch among ionic, covalent, and metallic types. It is a continuum. For two atoms A and B, a useful qualitative descriptor is the electronegativity contrast

$$\Delta\chi_{AB} = |\chi_A - \chi_B|. \tag{62}$$

Large $\Delta\chi_{AB}$ supports ionic character, while small contrast between metallic elements supports metallic bonding. Directional main-group bonds with appropriate distances support covalent character. Hydrogen bonds require a donor, hydrogen, acceptor, distance window, and angular geometry. The platform uses these ideas to present a chemistry-aware topology while preserving the important scientific caveat: distance and electronegativity alone do not determine a quantum-mechanical bond order in every material.

A professional platform should therefore distinguish between *assigned topology* and *electronic bonding*. Topology is the graph used for visualization, selection, reporting, and simulation preparation. Electronic bonding requires density, orbital, charge, ELF, or wavefunction-derived evidence. Aurora Nano keeps these views connected but not confused.

15.4 Coordination Polyhedra as Local Chemical Units

For a central atom i with neighbor set $\mathcal{N}(i)$, the local coordination polyhedron can be characterized by its centroid

$$\mathbf{c}_i = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{r}_j, \tag{63}$$

and by the covariance matrix of neighbor positions

$$C_i = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} (\mathbf{r}_j - \mathbf{c}_i)(\mathbf{r}_j - \mathbf{c}_i)^T. \tag{64}$$

The eigenvalues of C_i provide a compact description of anisotropy. Nearly equal eigenvalues indicate a roughly isotropic environment; a dominant eigenvalue indicates a linear or elongated coordination; two dominant eigenvalues indicate planar coordination. When combined with ECoN, bond-valence residuals, and angular histograms, these descriptors give a chemically meaningful view of local structure without claiming to replace expert interpretation.

15.5 Crystal-Field and Ligand-Environment Context

Transition-metal chemistry often requires a ligand-field viewpoint. In an ideal octahedral environment, the five d orbitals split into lower t_{2g} and higher e_g sets. A simple crystal-field stabilization estimate can be written as

$$E_{\mathrm{CFSE}} = \left(-0.4\, n_{t_{2g}} + 0.6\, n_{e_g}\right) \Delta_o + mP, \tag{65}$$

where Δ_o is the octahedral splitting, P is a pairing-energy term, $n_{t_{2g}}$ and n_{e_g} are orbital occupations, and m counts additional electron-pairing contributions in the chosen model. Aurora Nano does not need to expose proprietary electronic-structure heuristics to use this language professionally. It can report local coordination, symmetry, bond lengths, distortion, magnetic moments, and density-derived fields so that the user can interpret ligand environments within accepted chemical theory [58].

15.6 Charge, Spin, and Local Moment Consistency

For spin-polarized data, the spin density is

$$m(\mathbf{r}) = \rho_\uparrow(\mathbf{r}) - \rho_\downarrow(\mathbf{r}), \tag{66}$$

and the total charge density is

$$\rho(\mathbf{r}) = \rho_\uparrow(\mathbf{r}) + \rho_\downarrow(\mathbf{r}). \tag{67}$$

A local moment over region Ω_i is therefore

$$\mu_i = \int_{\Omega_i} m(\mathbf{r})\, d\mathbf{r}. \tag{68}$$

The professional reporting challenge is not only to plot $m(\mathbf{r})$, but to make clear which region Ω_i is implied: atomic sphere, Bader basin, Voronoi basin, grid mask, or user-defined selection. The platform can disclose this scientific distinction without disclosing internal field-processing optimizations.

16 Statistical Mechanics of Atomistic Observables

16.1 Ensemble Averages and Time Averages

Many observables in molecular simulation are ensemble averages. For a microscopic observable $A(\Gamma)$ over phase-space point Γ, the canonical expectation is

$$\langle A \rangle = \frac{1}{Z} \int A(\Gamma) \exp[-\beta H(\Gamma)] \, d\Gamma, \qquad Z = \int \exp[-\beta H(\Gamma)] \, d\Gamma, \tag{69}$$

where $\beta = (k_B T)^{-1}$ and H is the Hamiltonian. A trajectory analysis typically estimates this average from finite samples:

$$\overline{A} = \frac{1}{N_t} \sum_{m=1}^{N_t} A(\Gamma_m). \tag{70}$$

The difference between $\langle A \rangle$ and $\overline{A}$ is not a software detail. It is a scientific sampling issue. Aurora Nano therefore treats trajectory results as finite-time estimates whose quality depends on sampling length, equilibration, timestep, thermostat, barostat, and system size [12, 13, 62].

16.2 RDF Normalization and Coordination Shell Integrals

The radial distribution function in Eq. 16 can be interpreted as a density ratio:

$$g_{AB}(r) = \frac{\rho_B(r \mid A)}{\rho_B}, \tag{71}$$

where $\rho_B(r \mid A)$ is the conditional density of species B at distance r from species A. The expected number of B atoms in a shell $[r_1, r_2]$ around an A atom is

$$N_{AB}(r_1, r_2) = 4\pi \rho_B \int_{r_1}^{r_2} r^2 g_{AB}(r) \, dr. \tag{72}$$

This equation connects the plotted RDF to a chemically interpretable coordination number. In practice, the first coordination shell is often defined by integrating from $r = 0$ to the first minimum after the first peak. The platform can report this integral when the minimum is well defined and can flag cases where broad or noisy peaks make automatic shell interpretation unreliable.

16.3 Potential of Mean Force

For a pair of species in equilibrium, the potential of mean force is related to the RDF by

$$W_{AB}(r) = -k_B T \ln g_{AB}(r) + C, \tag{73}$$

where C is an arbitrary additive constant. Peaks in $g(r)$ correspond to local minima in $W(r)$; depleted regions correspond to barriers. This relationship is useful for liquids, electrolytes, polymer melts, and solvation environments. The scientific caveat is that $W(r)$ is a many-body free-energy projection, not a two-body potential suitable for direct reuse as a force field.

16.4 Static Structure Factor and Compressibility Limit

For isotropic systems, the total correlation function $h(r) = g(r) - 1$ connects to the structure factor through

$$S(q) = 1 + 4\pi\rho \int_0^\infty r^2 h(r) \frac{\sin(qr)}{qr}\, dr. \tag{74}$$

The small-q limit is related to the isothermal compressibility:

$$S(0) = \rho k_B T \kappa_T. \tag{75}$$

These equations help users understand why long-range sampling, box size, and low-q resolution matter. A visually smooth RDF is not automatically sufficient for accurate thermodynamic inference. A professional platform should expose enough physical context to prevent overinterpretation.

16.5 Diffusion Tensor and Anisotropic Transport

The scalar diffusion coefficient in Eq. 19 assumes isotropic transport. For crystals, layered materials, channels, and interfaces, the diffusion tensor is more appropriate:

$$D_{\alpha\beta} = \lim_{t\to\infty} \frac{1}{2t} \left\langle \Delta r_\alpha(t) \Delta r_\beta(t) \right\rangle. \tag{76}$$

The trace recovers the isotropic average:

$$D_{\text{iso}} = \frac{1}{3} \operatorname{tr}(D). \tag{77}$$

In an anisotropic solid electrolyte, D_{zz} may differ from D_{xx} and D_{yy} by orders of magnitude. Aurora Nano can therefore present species-resolved and direction-resolved mobility summaries when trajectory and cell information support the interpretation.

16.6 Green-Kubo Relations

The velocity autocorrelation function gives a second route to diffusion:

$$D = \frac{1}{3} \int_0^\infty \langle \mathbf{v}(0) \cdot \mathbf{v}(t) \rangle \, dt. \tag{78}$$

For ionic conductivity, the charge-current autocorrelation relation is

$$\sigma = \frac{1}{3Vk_BT} \int_0^\infty \langle \mathbf{J}(0) \cdot \mathbf{J}(t) \rangle \, dt, \qquad \mathbf{J}(t) = \sum_i q_i \mathbf{v}_i(t). \tag{79}$$

Equation 79 is powerful but sensitive to sampling, finite-size effects, charge models, and long-time noise [18]. Aurora Nano treats such calculations as heavy trajectory analyses that require careful execution and expert review.

16.7 Uncertainty of Time-Correlation Estimates

For a correlated trajectory, the effective number of independent samples is smaller than the number of frames. If τ_A is the integrated autocorrelation time of observable A, an approximate standard error scales as

$$\mathrm{SE}(\overline{A}) \approx \sqrt{\frac{2\tau_A}{T_{\mathrm{traj}}}}\,\sigma_A, \tag{80}$$

where T_{traj} is the total trajectory duration and σ_A is the standard deviation. The platform can use this principle in reporting language: a number obtained from a short trajectory should not be presented with false precision simply because many frames were loaded.

17 Defect, Strain, and Microstructure Physics Extensions

17.1 Least-Squares Origin of the Local Deformation Gradient

Section 6 introduced the Falk-Langer style deformation-gradient fit. The mathematical origin is a weighted least-squares problem over neighbor vectors:

$$\Pi(F) = \sum_j w_j \left\| \Delta\mathbf{r}_j^{\mathrm{cur}} - F\Delta\mathbf{r}_j^{\mathrm{ref}} \right\|^2. \tag{81}$$

Setting the derivative with respect to F to zero yields

$$\sum_j w_j \Delta\mathbf{r}_j^{\mathrm{cur}} \otimes \Delta\mathbf{r}_j^{\mathrm{ref}} = F \sum_j w_j \Delta\mathbf{r}_j^{\mathrm{ref}} \otimes \Delta\mathbf{r}_j^{\mathrm{ref}}, \tag{82}$$

and therefore

$$F = \left(\sum_j w_j \Delta\mathbf{r}_j^{\mathrm{cur}} \otimes \Delta\mathbf{r}_j^{\mathrm{ref}} \right) \left(\sum_j w_j \Delta\mathbf{r}_j^{\mathrm{ref}} \otimes \Delta\mathbf{r}_j^{\mathrm{ref}} \right)^{-1}. \tag{83}$$

The non-affine residual is the minimized value of Eq. 81. High residuals indicate local rearrangement not explained by a homogeneous deformation, which is why D^2_{min} is useful for shear transformation zones, amorphous plasticity, grain boundaries, and dislocation cores [20].

17.2 Polar Decomposition and Rotation-Strain Separation

The deformation gradient contains rotation and stretch. The polar decomposition separates them:

$$F = RU = VR, \tag{84}$$

where R is a proper rotation, U is the right stretch tensor, and V is the left stretch tensor. The right Cauchy-Green tensor is

$$C = F^T F = U^2, \tag{85}$$

and the Green-Lagrange strain in Eq. 56 follows as $E = (C - I)/2$. This decomposition matters visually. A rotated but undistorted crystal should not be interpreted as strained. A professional analysis platform must separate orientation from deformation when presenting local strain maps.

17.3 Scalar Strain Invariants

A tensor field is difficult to inspect atom by atom, so scalar invariants are often used. The hydrostatic strain component is

$$\varepsilon_{\mathrm{hyd}} = \frac{1}{3}\,\mathrm{tr}(\varepsilon), \tag{86}$$

and the deviatoric strain tensor is

$$\varepsilon' = \varepsilon - \varepsilon_{\mathrm{hyd}} I. \tag{87}$$

A von-Mises equivalent shear strain can be written as

$$\varepsilon_{\mathrm{VM}} = \sqrt{\frac{2}{3}\,\varepsilon' : \varepsilon'}. \tag{88}$$

These quantities support a physically interpretable color map: hydrostatic strain highlights dilation and compression, while deviatoric strain highlights shear-dominated distortion.

17.4 Centrosymmetry as a Broken-Inversion Measure

The centrosymmetry parameter in Section 6 can be interpreted as a local inversion-symmetry error. If an atom has n neighbors arranged in opposite pairs, each pair ideally satisfies

$$\mathbf{r}_{i_k} + \mathbf{r}_{j_k} = \mathbf{0}. \tag{89}$$

The CSP sums the squared deviations from this condition. Thermal vibrations produce a finite background, surfaces remove opposite partners, and defects break local symmetry more strongly. The platform can therefore present CSP as a defect-sensitive scalar, while making clear that temperature, surface proximity, and neighbor choice influence the numerical threshold [19].

17.5 Bond-Orientational Order and Local Symmetry

Steinhardt order parameters describe the angular distribution of neighbor bonds using spherical harmonics [24]. For atom i,

$$q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_{lm}(\hat{\mathbf{r}}_{ij}), \tag{90}$$

and the rotationally invariant order parameter is

$$q_l(i) = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |q_{lm}(i)|^2 \right)^{1/2}. \tag{91}$$

Different local structures occupy different regions of the (q_4, q_6) descriptor space. FCC, HCP, BCC, icosahedral, liquid-like, and amorphous environments can often be separated when noise and temperature are controlled. Aurora Nano can combine these descriptors with CNA, PTM, and CSP to create a multi-evidence local-structure report rather than relying on one metric alone.

17.6 Misorientation and Grain Context

For two local orientation matrices R_a and R_b, the relative rotation is

$$\Delta R = R_b R_a^T. \tag{92}$$

The misorientation angle is

$$\theta = \cos^{-1}\left(\frac{\mathrm{tr}(\Delta R) - 1}{2} \right). \tag{93}$$

Crystal symmetry reduces the physically distinct orientation space, so professional grain analysis reports should state the symmetry convention used. Aurora Nano can expose misorientation, cluster labels, twin candidates, and boundary indicators as reviewable scientific artifacts while keeping production segmentation logic behind the platform boundary.

18 Volumetric Field Theory and Electronic Observables

18.1 Periodic Grid Integrals

A periodic scalar field $\rho(\mathbf{r})$ sampled on a grid $N_x \times N_y \times N_z$ represents values at fractional grid points. The grid-cell volume element is

$$\Delta V = \frac{V}{N_x N_y N_z}. \tag{94}$$

The integral of the field over the cell is approximated by

$$Q = \int_\Omega \rho(\mathbf{r})\, d\mathbf{r} \approx \Delta V \sum_{u=1}^{N_x} \sum_{v=1}^{N_y} \sum_{w=1}^{N_z} \rho_{uvw}. \tag{95}$$

For an electron density, Q should correspond to the electron count represented by the calculation and file convention. For a spin density, the same integral gives a magnetic moment in units implied by the source. This is why volumetric import cannot be treated as only a graphics problem: units, grid vectors, origin, periodicity, and normalization all have scientific meaning.

18.2 Charge-Difference Density

A charge-difference density is commonly used to inspect bonding, charge transfer, adsorption, and interfacial redistribution:

$$\Delta\rho(\mathbf{r}) = \rho_{AB}(\mathbf{r}) - \rho_A(\mathbf{r}) - \rho_B(\mathbf{r}). \tag{96}$$

Positive regions indicate accumulation relative to the chosen references; negative regions indicate depletion. The interpretation depends strongly on the reference states A and B, their geometries, and alignment on the same grid. Aurora Nano can visualize and integrate $\Delta\rho$, but the paper should make clear that physical conclusions require consistent reference calculations.

18.3 Gradient, Laplacian, and Critical-Point Context

The gradient of a scalar field is

$$\nabla\rho = \left(\frac{\partial\rho}{\partial x}, \frac{\partial\rho}{\partial y}, \frac{\partial\rho}{\partial z}\right), \tag{97}$$

and the Laplacian is

$$\nabla^2\rho = \frac{\partial^2\rho}{\partial x^2} + \frac{\partial^2\rho}{\partial y^2} + \frac{\partial^2\rho}{\partial z^2}. \tag{98}$$

In quantum-topological analysis, critical points satisfy $\nabla\rho = 0$, and the Hessian eigenvalues classify local maxima, minima, and saddle points [31]. Aurora Nano's volumetric viewport and slice panels are designed to support visual exploration of these fields. Formal topological classification can be routed to specialized backends when the analysis requires production-grade basin integration.

18.4 Electron Localization Function

The electron localization function (ELF) compares local kinetic-energy density behavior to that of a homogeneous electron gas [30]. In one common spin-unpolarized form,

$$\mathrm{ELF}(\mathbf{r}) = \frac{1}{1 + [D(\mathbf{r})/D_h(\mathbf{r})]^2}, \tag{99}$$

where

$$D(\mathbf{r}) = \tau(\mathbf{r}) - \frac{1}{4}\frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})}, \tag{100}$$

and

$$D_h(\mathbf{r}) = \frac{3}{5}(6\pi^2)^{2/3}\rho(\mathbf{r})^{5/3}. \tag{101}$$

ELF values near one often indicate localized electron pairs, lone pairs, or covalent bonding regions; lower values indicate more delocalized behavior. The platform should present ELF as an interpretive field, not as

a standalone bond-definition oracle. Chemistry emerges from the combined evidence of geometry, topology, density, spin, and energetic context.

18.5 Poisson Relation and Electrostatic Potential

The electrostatic potential $\phi(\mathbf{r})$ is connected to charge density by Poisson's equation:

$$\nabla^2\phi(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0}. \tag{102}$$

For periodic solids, the zero of potential is convention-dependent, and charged cells require careful treatment. Aurora Nano can visualize potential fields and slices while preserving this scientific limitation in reporting language. A potential color map without unit, reference, and boundary-condition context is not a complete scientific result.

18.6 Isosurfaces as Level Sets

An isosurface is the level set

$$\mathcal{S}_c = \{\mathbf{r} : \rho(\mathbf{r}) = c\}\,. \tag{103}$$

The chosen value c is part of the scientific statement. A low isovalue may show diffuse features and connections; a high isovalue may isolate local maxima. For spin density, opposite signs should be rendered distinctly. For charge-difference density, symmetric positive and negative isovalues support balanced interpretation. Aurora Nano's rendering surface therefore acts as a scientific visualization of a defined mathematical object, not as a decorative mesh.

19 Diffraction and Reciprocal-Space Reporting Extensions

19.1 Scattering Vector and Bragg Condition

For incident and scattered wavevectors $\mathbf{k}_i$ and $\mathbf{k}_s$ with $|\mathbf{k}_i| = |\mathbf{k}_s| = 2\pi/\lambda$, the scattering vector is

$$\mathbf{Q} = \mathbf{k}_s - \mathbf{k}_i. \tag{104}$$

Elastic diffraction occurs when $\mathbf{Q}$ equals a reciprocal-lattice vector $\mathbf{g}_{hkl}$. Combining Eq. 51 with diffraction geometry gives Bragg's law:

$$2d_{hkl}\sin\theta = \lambda. \tag{105}$$

This chain of equations connects the user-visible powder plot directly to the unit cell, reciprocal lattice, and wavelength. It also provides a diagnostic path: a shifted peak may indicate lattice-parameter error, wavelength mismatch, strain, instrument zero offset, or composition-dependent expansion.

19.2 Structure Factors and Thermal Displacement

For a reflection (hkl), the structure factor is

$$F_{hkl} = \sum_{j=1}^{N} f_j(Q)\, o_j\, \exp\left[2\pi i(hx_j + ky_j + lz_j)\right] \exp(-B_j s^2), \tag{106}$$

where $f_j(Q)$ is the atomic scattering factor, o_j is occupancy, (x_j, y_j, z_j) are fractional coordinates, B_j is an isotropic displacement parameter, and $s = \sin\theta/\lambda$. The ideal integrated intensity is proportional to

$$I_{hkl} \propto m_{hkl} |F_{hkl}|^2, \tag{107}$$

where m_{hkl} is reflection multiplicity. In a professional report, this equation helps the user understand why missing atoms, wrong occupancies, thermal disorder, and symmetry choices change relative peak intensities even when peak positions are correct [63].

19.3 Profile Functions and Residual Metrics

A powder pattern compares observed intensities y_i^{obs} to calculated intensities y_i^{calc}. The weighted profile residual is

$$R_{wp} = \left[\frac{\sum_i w_i \left(y_i^{\mathrm{obs}} - y_i^{\mathrm{calc}}\right)^2}{\sum_i w_i \left(y_i^{\mathrm{obs}}\right)^2} \right]^{1/2}. \tag{108}$$

The expected residual is

$$R_{\mathrm{exp}} = \left[\frac{N - P}{\sum_i w_i \left(y_i^{\mathrm{obs}}\right)^2} \right]^{1/2}, \tag{109}$$

where N is the number of observations and P is the number of refined parameters. The goodness of fit is

$$\chi^2 = \left(\frac{R_{wp}}{R_{\mathrm{exp}}} \right)^2. \tag{110}$$

Aurora Nano can display these quantities for Rietveld-style review while being clear that automated profile matching does not replace crystallographic refinement expertise [44, 47].

19.4 Size-Strain Broadening

Peak broadening can come from crystallite size, microstrain, instrument resolution, stacking faults, or sample preparation. A common first diagnostic is the Williamson-Hall form

$$\beta \cos\theta = \frac{K\lambda}{L} + 4\epsilon \sin\theta, \tag{111}$$

where β is the peak width after instrument correction, K is a shape factor, L is an apparent coherence length, and ϵ is a microstrain estimate. This equation is useful for reporting context, but Aurora Nano should avoid overstating it. A single broadened pattern can rarely identify a unique physical cause without instrument calibration and sample knowledge.

19.5 Patterson and Pair-Vector Interpretation

The Patterson map in Eq. 43 is the autocorrelation of electron density. In continuous form,

$$P(\mathbf{u}) = \int \rho(\mathbf{r})\rho(\mathbf{r}+\mathbf{u})\, d\mathbf{r}. \tag{112}$$

This makes the interpretation intuitive: peaks correspond to common interatomic vectors, weighted by scattering power. Aurora Nano can render Patterson slices and peaks to help crystallographers inspect heavy-atom substructures, but the platform should present it as a guide for expert analysis rather than an automatic structure solution.

20 Thermodynamic and Materials-Discovery Derivations

20.1 Formation Energy and Reference States

For a compound A_xB_y, the formation energy per atom is

$$\Delta E_f(A_xB_y) = \frac{E(A_xB_y) - x\mu_A^0 - y\mu_B^0}{x+y}, \tag{113}$$

where μ_A^0 and μ_B^0 are reference chemical potentials of the elemental phases. In a multicomponent system,

$$\Delta E_f = \frac{E(\{n_i\}) - \sum_i n_i \mu_i^0}{\sum_i n_i}. \tag{114}$$

The choice of reference states is not a software implementation detail; it is a thermodynamic convention. Aurora Nano reports phase-stability quantities in a way that keeps reference energies, composition normalization, and data source visible to the reviewer.

20.2 Convex Hull and Energy Above Hull

Let $\mathbf{x}$ denote a composition vector and $E(\mathbf{x})$ the formation energy. The lower convex envelope $E_{\text{hull}}(\mathbf{x})$ is the minimum energy obtainable by phase separation into available competing phases. The energy above hull is

$$E_{\text{above}}(\mathbf{x}) = E(\mathbf{x}) - E_{\text{hull}}(\mathbf{x}). \tag{115}$$

A stable phase has $E_{\text{above}} = 0$ within numerical tolerance. A metastable phase has $E_{\text{above}} > 0$ but may still be experimentally accessible because of kinetics, temperature, pressure, entropy, defects, particle size, or synthesis pathway. This distinction is important for a professional materials-discovery platform: the hull is a thermodynamic screen, not a final experimental verdict [42, 43].

20.3 Chemical-Potential Constraints

For a compound $A_xB_yC_z$, equilibrium with its constituents imposes

$$x\mu_A + y\mu_B + z\mu_C = \Delta H_f(A_xB_yC_z). \tag{116}$$

Avoiding precipitation of a competing phase $A_pB_qC_r$ requires

$$p\mu_A + q\mu_B + r\mu_C \leq \Delta H_f(A_pB_qC_r). \tag{117}$$

The feasible chemical-potential region is therefore a polytope defined by linear inequalities. Aurora Nano can visualize this region as a decision-support map for growth windows, defect chemistry, and interface stability. The mathematical model is public; the product-specific data integration and ranking policies remain within the commercial boundary.

20.4 Surface and Interface Energies

For a symmetric slab with two equivalent surfaces of area A, the surface energy is

$$\gamma = \frac{E_{\text{slab}} - NE_{\text{bulk}}}{2A}. \tag{118}$$

For an interface between materials 1 and 2, an interface energy can be expressed as

$$\gamma_{12} = \frac{E_{12} - N_1E_1 - N_2E_2}{A}, \tag{119}$$

where E_{12} is the total energy of the combined interface model. These expressions are simple, but the physical interpretation depends on stoichiometry, chemical potentials, strain, slab thickness, vacuum thickness, dipole corrections, and relaxation protocol. Aurora Nano can expose the formula and metadata while keeping proprietary interface-matching and scoring heuristics confidential.

20.5 Adsorption Energy

For a molecule or atom adsorbed on a surface, the adsorption energy is commonly written as

$$E_{\text{ads}} = E_{\text{surface+adsorbate}} - E_{\text{surface}} - E_{\text{adsorbate}}. \tag{120}$$

With this sign convention, negative values indicate exothermic adsorption. The platform should always state the sign convention, reference state, charge state, spin state, coverage, and cell size. A professional product can assist the user by organizing these quantities without implying that one number alone fully describes catalytic behavior.

20.6 Grand Potential for Open Systems

When a system exchanges particles with reservoirs, the grand potential is

$$\Phi = E - \sum_i \mu_i N_i. \tag{121}$$

Grand-potential diagrams are useful for defects, surfaces, interfaces, and phase stability under growth conditions. The same structural model can be stable or unstable depending on the chemical-potential environment. Aurora Nano's role is to make those dependencies explicit and traceable, not to hide them behind a single unqualified stability label.

21 Enterprise Scientific Platform Assurances

21.1 From Visualization to Governed Scientific Workflows

The enterprise value of Aurora Nano Studio comes from joining scientific analysis with operational governance. In a conventional desktop workflow, the researcher may load a file, make a picture, export a screenshot, run a separate script, and manually record what happened. Aurora Nano treats the same activity as a governed analysis process: a structure, trajectory, volume, diffraction trace, or compute result becomes an asset; the analysis becomes a typed operation; the output becomes a reportable artifact; and the path from input to result remains traceable.

This does not require exposing proprietary internals. The reader only needs to understand the professional guarantees: reproducible modifier specifications, explicit compute targets, auditable job state, large-artifact policy, and validated scientific output definitions.

21.2 Lineage as a Scientific Object

For an analysis result R, a simplified lineage record can be written as

$$\mathcal{L}(R) = (S,\, A,\, \Theta,\, T,\, P,\, H)\,, \tag{122}$$

where S is the source asset, A is the analysis operation, Θ is the parameter set, T is the compute target, P is the platform version or wire contract, and H is a content or result hash. The exact implementation of H, target selection, and persistence policy can remain proprietary. The important professional statement is that a result is not detached from its origin.

21.3 Governance Without Scientific Friction

Governance need	Scientific risk if absent	Aurora Nano platform posture
Asset provenance	Users cannot determine which file, frame, or grid produced a plot.	Source assets and derived artifacts are linked through metadata and job records.
Parameter traceability	A reported value cannot be reproduced because the cutoff, bin count, species filter, or frame range is unknown.	Modifier and analysis specifications are stored as structured parameters.
Compute target clarity	Heavy jobs may silently run in unsuitable environments or produce inconsistent results.	The compute boundary is explicit: local, worker, remote, or cloud target.
Large-artifact policy	Sensitive or massive simulation files may be loaded or uploaded without appropriate review.	Manifest-only and explicit-upload policies separate metadata indexing from content transfer.
Auditability	Reports become screenshots without operational evidence.	Job state, events, exports, and activity records support review.

This framing is especially important for industrial users in pharma, semiconductor, aerospace, energy, and advanced materials. In those environments, the question is not only whether the visualization is attractive. The question is whether the analysis can be repeated, governed, explained, and defended.

21.4 Compute Routing as a Product Boundary

Aurora Nano's public story should emphasize that different workloads have different execution needs. A small geometric inspection can remain interactive. A large trajectory, high-resolution volumetric operation, or heavy statistical analysis may require a worker or remote scientific backend. The white paper can disclose the existence of this compute boundary without documenting exact scheduling policy.

$$\text{operation} + \text{data size} + \text{policy} + \text{availability} \longrightarrow \text{approved compute target}. \tag{123}$$

Equation 123 is intentionally conceptual. It communicates platform maturity while preserving the proprietary target resolver, cost model, and operational tuning.

21.5 Confidentiality Statement for Client-Facing Use

The following language can be used in executive or technical-review settings:

> Aurora Nano Studio documents the mathematical definitions, physical interpretation, validation scope, and enterprise governance model of its analysis outputs. Certain production details—including internal routing heuristics, performance-specific data layouts, private validation corpora, partner-specific calibration policies, and commercial workflow controls—are proprietary to Aurora Simulations and are intentionally not disclosed. This boundary does not limit scientific review of the reported observables; it protects the product implementation while preserving transparent interpretation of results.

This statement directly addresses the product concern. It assures the reader that the platform is professional without converting the paper into a blueprint.

21.6 Quality Language for Reports

A scientific report generated by Aurora Nano should be readable by a domain expert. Each major result should identify:

- the source structure, trajectory, or field;
- the mathematical definition of the reported quantity;
- the physically meaningful units and normalization;
- the parameters that materially affect interpretation;
- the validation or limitation relevant to the operation;
- the lineage connection to the compute job or local analysis step.

This report structure communicates professionalism without exposing internal code. It also helps prevent misuse of advanced scientific visualizations in contexts where sampling, model choice, or domain limitations matter.

21.7 Scientific Review Checklist for Prospective Clients

The following checklist is intended for client-facing scientific review. It summarizes what a professional reader should be able to verify without receiving proprietary implementation details.

Review question	Evidence the platform should provide	Proprietary boundary preserved
What exactly was analyzed?	Source asset name, format, frame or model index, cell parameters, atom count, species list, and relevant metadata.	Internal file-ranking, indexing optimization, and ingestion heuristics.
Which scientific definition was used?	Equation-level observable definition, unit convention, normalization, and assumptions.	Private kernel implementation and performance-specific numerical layout.
Which parameters mattered?	Cutoffs, bin counts, species filters, isovalues, wavelength, temperature, frame range, and selected reference states.	Internal defaults governance, partner-specific presets, and ranking rules.
Can the result be reproduced?	Stored modifier specification, compute target, operation version, result artifact, and lineage tuple.	Private cache strategy, target scoring, and operational recovery policy.
What are the known limitations?	Scope statements for RDF periodicity, MVP-grade classifiers, force-field coverage, heavy trajectory execution, and large-artifact handling.	Internal roadmap sequencing and commercial validation gates.
How should the result be interpreted?	Physical meaning, uncertainty language, and cross-checks against complementary observables.	Proprietary report composition logic and user-specific review workflows.

This checklist is intentionally non-procedural. It does not teach the reader how to reproduce Aurora Nano's internal workflow. It teaches the reader how to trust the output as a governed scientific result. That distinction is central to positioning Aurora Nano as a professional platform rather than a collection of isolated analysis scripts.

22 Validation and Performance

22.1 Performance Benchmarks

Aurora Nano's core analysis paths have been validated against the following timing budgets on current commodity hardware:

Workload	Verified budget
10 000-atom bounding-box plus nearest-neighbour scan	< 800 ms
100 000-atom bounding-box scan	< 1500 ms
64^3 scalar grid upsampled to 128^3 subset	< 1500 ms
64-atom model electron density on 40^3 periodic grid	< 1500 ms

Performance-sensitive paths use contiguous typed-array storage. Bounding volume hierarchy (BVH) nodes

are stored as flat `Float64Array` records with `Uint32Array` primitive order [57]. Volume fields are stored as `Float32Array`. Trajectory memory is estimated as three 64-bit components per atom per frame.

22.2 Test Coverage

The test suite covers parsers (XYZ, POSCAR, CIF, mmCIF, LAMMPS, GRO, MOL, MOL2, PDB, VTK, CUBE, XSF, CHGCAR), crystallographic mathematics (cell matrix, lattice transformations, PBC distance, coordination), modifier behaviour (RDF, ADF, MSD, VACF, centrosymmetry, PTM, CNA, atomic strain, cluster analysis, ring finding, bond-valence), distribution normalisation, phase-hull construction, chemical-potential polygon computation, electronic structure parsing, and import-policy guards for large files. Parser round-trips are verified to reproduce atom counts, cell parameters, and fractional coordinates within numerical tolerance.

22.3 Stated Limitations

Aurora Nano is designed for transparency about its current boundaries:

- RDF applies the minimum-image convention for orthogonal cells; non-orthogonal RDF uses raw Cartesian distances in the current implementation. Full lattice-vector PBC is used for bond and coordination analyses.
- PTM, grain segmentation, dislocation extraction, and Voronoi topology are labelled as MVP entries in the modifier registry; robustness across all system types and temperatures has not been fully validated.
- Force-field assignment and simulation deck generation require correct force-field coverage of the atom types present; missing parameters are reported but not automatically resolved.
- All reports and scientific plots are intended to support expert scientific review. They do not replace domain peer review, instrument calibration, or certified metrology procedures.
- Heavy trajectory analysis (Van Hove, dielectric, ionic conductivity, VDOS from velocity) requires server-side execution; results are not available in offline-only mode.
- Large simulation artifacts (>10 MB) trigger an import-review step that estimates atom count and parse cost before loading into memory.

23 Applications

Domain	Representative Tasks	Relevant Sections
Materials science	Local-structure classification in metals and alloys; defect identification; phase-stability screening; RDF and PDF comparison.	§2, §4, §6, §9
Computational chemistry	SMILES ingestion; polymer chain construction; force-field typing; simulation deck generation; molecular geometry inspection.	§3, §12

Domain	Representative Tasks	Relevant Sections
Condensed-matter physics	MSD/VACF/VDOS analysis; Lindemann index; structure-factor computation; electronic band structure and DOS visualisation; Fermi-surface estimation.	§5, §4, §11
Crystallography	Space-group symmetry; Wyckoff expansion; Niggli reduction; powder pattern simulation; Rietveld/Pawley/Le Bail comparison; Patterson maps; morphology.	§2, §10
Semiconductor engineering	Heterointerface lattice matching; mismatch strain; adhesion energy; slab preparation; DFT input generation; band-gap determination.	§8, §11
Catalysis	Adsorption site identification; DFT adsorption-energy attachment; surface slab construction; passivation; bond-valence diagnostics.	§3, §8
Battery and energy materials	Phase-stability diagrams; chemical-potential windows; ion-diffusion MSD; charge-density visualisation; ELF bonding analysis.	§9, §5, §7
Pharmaceuticals and soft matter	SMILES-to-structure; polymer packing; force-field assignment; H-bond analysis; VACF/VDOS for amorphous materials; Lindemann disorder.	§3, §12, §5
Aerospace and structural alloys	Grain-boundary analysis; slip-vector dislocation maps; elastic strain; twin-domain visualisation; centrosymmetry defect maps.	§6

24 Conclusion

Aurora Nano Studio consolidates a broad set of atomistic analysis methods into a single interactive environment. The implementation covers crystallographic foundations, chemistry-aware bonding and molecular topology, radial and angular distribution functions, dynamical trajectory analysis, a comprehensive suite of defect and local-structure classifiers, periodic volumetric field analysis, surface and interface preparation, thermodynamic phase stability, powder diffraction simulation and profile analysis, electronic structure visualisation, and molecular simulation preparation—each grounded in the established scientific literature and accompanied by validated test coverage.

The environment is designed for both interactive scientific exploration and systematic programmatic workflows. Modifier pipelines are deterministic and portable, enabling reproducibility of analysis results across operators and compute environments. Large simulation artifacts are handled through a policy-based import system that avoids silent memory failures. Heavy analyses that exceed interactive budgets are routed to server-side scientific Python workers, returning typed result artifacts that integrate with the same visualisation panels used for browser-side results.

The methods described in this white paper reflect the implemented capability as of July 2026. Known limitations, particularly around non-orthogonal RDF, MVP-grade defect classifiers, and heavy-trajectory-analysis server dependencies, are documented explicitly. Continued development priorities include broader force-field coverage, production-grade dislocation extraction, full non-orthogonal PBC for all distribution functions, and extended Fermi-surface analysis for complex metallic systems.

References

[1] Hahn, T. (ed.) *International Tables for Crystallography, Volume A: Space-Group Symmetry*, 5th ed. Springer, Dordrecht, 2006.

[2] Krivy, I. and Gruber, B. "A unified algorithm for determining the reduced Niggli cell." *Acta Crystallographica Section A*, 32, 1976, 297–298. DOI: 10.1107/S0567739476000636.

[3] Grosse-Kunstleve, R. W., Sauter, N. K., and Adams, P. D. "Numerically stable algorithms for the computation of reduced unit cells." *Acta Crystallographica Section A*, 60, 2004, 1–6. DOI: 10.1107/S010876730302186X.

[4] Cordero, B. et al. "Covalent radii revisited." *Dalton Transactions*, 2008, 2832–2838. DOI: 10.1039/B801115J.

[5] Allen, F. H. et al. "Tables of bond lengths determined by X-ray and neutron diffraction." *Journal of the Chemical Society, Perkin Transactions 2*, 1987, S1–S19.

[6] Brown, I. D. and Altermatt, D. "Bond-valence parameters obtained from a systematic analysis of the Inorganic Crystal Structure Database." *Acta Crystallographica Section B*, 41, 1985, 244–247. DOI: 10.1107/S0108768185002063.

[7] Ewald, P. P. "Die Berechnung optischer und elektrostatischer Gitterpotentiale." *Annalen der Physik*, 369, 1921, 253–287. DOI: 10.1002/andp.19213690304.

[8] Hoppe, R. "Effective coordination numbers (ECoN) and mean fictive ionic radii (MEFIR)." *Zeitschrift für Kristallographie*, 150, 1979, 23–52.

[9] Weininger, D. "SMILES, a chemical language and information system." *Journal of Chemical Information and Computer Sciences*, 28, 1988, 31–36. DOI: 10.1021/ci00057a005.

[10] RDKit: Open-Source Cheminformatics. https://www.rdkit.org. Accessed July 2026.

[11] Blaney, J. M. and Dixon, J. S. "Distance geometry in molecular modeling." *Reviews in Computational Chemistry*, 5, 1994, 299–335. DOI: 10.1002/9780470125823.ch6.

[12] Allen, M. P. and Tildesley, D. J. *Computer Simulation of Liquids*. Oxford University Press, Oxford, 1987.

[13] Frenkel, D. and Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. Academic Press, San Diego, 2002.

[14] Hansen, J.-P. and McDonald, I. R. *Theory of Simple Liquids*, 3rd ed. Academic Press, Amsterdam, 2006.

[15] Debye, P. "Zerstreuung von Röntgenstrahlen." *Annalen der Physik*, 351, 1915, 809–823. DOI: 10.1002/andp.19153510606.

[16] Einstein, A. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen." *Annalen der Physik*, 322, 1905, 549–560. DOI: 10.1002/andp.19053220806.

[17] Lindemann, F. A. "Über die Berechnung molekularer Eigenfrequenzen." *Physikalische Zeitschrift*, 11, 1910, 609–612.

[18] Kubo, R. "The fluctuation-dissipation theorem." *Reports on Progress in Physics*, 29, 1966, 255–284. DOI: 10.1088/0034-4885/29/1/306.

[19] Kelchner, C. L., Plimpton, S. J., and Hamilton, J. C. "Dislocation nucleation and defect structure during surface indentation." *Physical Review B*, 58, 1998, 11085–11088. DOI: 10.1103/PhysRevB.58.11085.

[20] Falk, M. L. and Langer, J. S. "Dynamics of viscoplastic deformation in amorphous solids." *Physical Review E*, 57, 1998, 7192–7205. DOI: 10.1103/PhysRevE.57.7192.

[21] Faken, D. and Jónsson, H. "Systematic analysis of local atomic structure combined with 3D computer graphics." *Computational Materials Science*, 2, 1994, 279–286. DOI: 10.1016/0927-0256(94)90109-0.

[22] Ackland, G. J. and Jones, A. P. "Applications of local crystal structure measures in experiment and simulation." *Physical Review B*, 73, 2006, 054104. DOI: 10.1103/PhysRevB.73.054104.

[23] Larsen, P. M., Schmidt, S., and Schiøtz, J. "Robust structural identification via polyhedral template matching." *Modelling and Simulation in Materials Science and Engineering*, 24, 2016, 055007. DOI: 10.1088/0965-0393/24/5/055007.

[24] Steinhardt, P. J., Nelson, D. R., and Ronchetti, M. "Bond-orientational order in liquids and glasses." *Physical Review B*, 28, 1983, 784–805. DOI: 10.1103/PhysRevB.28.784.

[25] Molinero, V. and Moore, E. B. "Water modeled as an intermediate element between carbon and silicon." *Journal of Physical Chemistry B*, 113, 2009, 4008–4016. DOI: 10.1021/jp805227c.

[26] Zimmerman, J. A., Kelchner, C. L., Klein, P. A., Hamilton, J. C., and Foiles, S. M. "Surface step effects on nanoindentation." *Physical Review Letters*, 87, 2001, 165507. DOI: 10.1103/PhysRevLett.87.165507.

[27] Voronoi, G. "Nouvelles applications des paramètres continus à la théorie des formes quadratiques." *Journal für die reine und angewandte Mathematik*, 133, 1908, 97–178. DOI: 10.1515/crll.1908.133.97.

[28] Rycroft, C. H. "VORO++: A three-dimensional Voronoi cell library in C++." *Chaos*, 19, 2009, 041111. DOI: 10.1063/1.3215722.

[29] Tarjan, R. E. "Efficiency of a good but not linear set union algorithm." *Journal of the ACM*, 22, 1975, 215–225. DOI: 10.1145/321879.321884.

[30] Becke, A. D. and Edgecombe, K. E. "A simple measure of electron localization in atomic and molecular systems." *Journal of Chemical Physics*, 92, 1990, 5397–5403. DOI: 10.1063/1.458517.

[31] Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*. Oxford University Press, Oxford, 1990.

[32] Koch, W. and Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, 2nd ed. Wiley-VCH, Weinheim, 2001.

[33] Lorensen, W. E. and Cline, H. E. "Marching Cubes: A high resolution 3D surface construction algorithm." *Computer Graphics*, 21, 1987, 163–169. DOI: 10.1145/37402.37422.

[34] Zur, A. and McGill, T. C. "Lattice match: an application to heteroepitaxy." *Journal of Applied Physics*, 55, 1984, 378–386. DOI: 10.1063/1.333084.

[35] Lennard-Jones, J. E. "Cohesion." *Proceedings of the Physical Society*, 43, 1931, 461–482. DOI: 10.1088/0959-5309/43/5/301.

[36] Kresse, G. and Furthmüller, J. "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set." *Physical Review B*, 54, 1996, 11169–11186. DOI: 10.1103/PhysRevB.54.11169.

[37] Giannozzi, P. et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." *Journal of Physics: Condensed Matter*, 21, 2009, 395502. DOI: 10.1088/0953-8984/21/39/395502.

[38] Blöchl, P. E. "Projector augmented-wave method." *Physical Review B*, 50, 1994, 17953–17979. DOI: 10.1103/PhysRevB.50.17953.

[39] Heyd, J., Scuseria, G. E., and Ernzerhof, M. "Hybrid functionals based on a screened Coulomb potential." *Journal of Chemical Physics*, 118, 2003, 8207–8215. DOI: 10.1063/1.1564060.

[40] Setyawan, W. and Curtarolo, S. "High-throughput electronic band structure calculations: Challenges and tools." *Computational Materials Science*, 49, 2010, 299–312. DOI: 10.1016/j.commatsci.2010.05.010.

[41] de Fontaine, D. "Cluster approach to order-disorder transformations in alloys." *Solid State Physics*, 47, 1994, 33–176. DOI: 10.1016/S0081-1947(08)60639-6.

[42] Ong, S. P. et al. "Python Materials Genomics (pymatgen): A robust, open-source Python library for materials analysis." *Computational Materials Science*, 68, 2013, 314–319. DOI: 10.1016/j.commatsci.2012.10.028.

[43] Bartel, C. J. et al. "A critical examination of compound stability predictions from machine-learned formation energies." *npj Computational Materials*, 6, 2020, 97. DOI: 10.1038/s41524-020-00362-y.

[44] Rietveld, H. M. "A profile refinement method for nuclear and magnetic structures." *Journal of Applied Crystallography*, 2, 1969, 65–71. DOI: 10.1107/S0021889869006558.

[45] Pawley, G. S. "Unit-cell refinement from powder diffraction scans." *Journal of Applied Crystallography*, 14, 1981, 357–361. DOI: 10.1107/S0021889881009618.

[46] Le Bail, A., Duroy, H., and Fourquet, J. L. "Ab-initio structure determination of $LiSbWO_6$ by X-ray powder diffraction." *Materials Research Bulletin*, 23, 1988, 447–452. DOI: 10.1016/0025-5408(88)90019-0.

[47] Thompson, P., Cox, D. E., and Hastings, J. B. "Rietveld refinement of Debye–Scherrer synchrotron X-ray data from Al_2O_3." *Journal of Applied Crystallography*, 20, 1987, 79–83. DOI: 10.1107/S0021889887087090.

[48] Donnay, J. D. H. and Harker, D. "A new law of crystal morphology extending the law of Bravais." *American Mineralogist*, 22, 1937, 446–467.

[49] Patterson, A. L. “A Fourier series method for the determination of the components of interatomic distances in crystals.” *Physical Review*, 46, 1934, 372–376. DOI: 10.1103/PhysRev.46.372.

[50] Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. “Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids.” *Journal of the American Chemical Society*, 118, 1996, 11225–11236. DOI: 10.1021/ja9621760.

[51] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. “Development and testing of a general Amber force field.” *Journal of Computational Chemistry*, 25, 2004, 1157–1174. DOI: 10.1002/jcc.20035.

[52] Qiu, Y. et al. “Development and benchmarking of open force field v1.0.0—the Parsley small molecule force field.” *Journal of Chemical Theory and Computation*, 17, 2021, 6262–6280. DOI: 10.1021/acs.jctc.1c00571.

[53] Eastman, P. et al. “OpenMM 7: Rapid development of high performance algorithms for molecular dynamics.” *PLOS Computational Biology*, 13, 2017, e1005659. DOI: 10.1371/journal.pcbi.1005659.

[54] Martínez, L., Andrade, R., Birgin, E. G., and Martínez, J. M. “PACKMOL: A package for building initial configurations for molecular dynamics simulations.” *Journal of Computational Chemistry*, 30, 2009, 2157–2164. DOI: 10.1002/jcc.21224.

[55] Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., and Beckstein, O. “MDAnalysis: A toolkit for the analysis of molecular dynamics simulations.” *Journal of Computational Chemistry*, 32, 2011, 2319–2327. DOI: 10.1002/jcc.21787.

[56] McGibbon, R. T. et al. “MDTraj: A modern open library for the analysis of molecular dynamics trajectories.” *Biophysical Journal*, 109, 2015, 1528–1532. DOI: 10.1016/j.bpj.2015.08.015.

[57] Wald, I., Boulos, S., and Shirley, P. “Ray tracing deformable scenes using dynamic bounding volume hierarchies.” *ACM Transactions on Graphics*, 26, 2007, 6. DOI: 10.1145/1189762.1206075.

[58] Pauling, L. *The Nature of the Chemical Bond*, 3rd ed. Cornell University Press, Ithaca, 1960.

[59] Born, M. and Huang, K. *Dynamical Theory of Crystal Lattices*. Oxford University Press, Oxford, 1954.

[60] Ashcroft, N. W. and Mermin, N. D. *Solid State Physics*. Holt, Rinehart and Winston, New York, 1976.

[61] Nye, J. F. *Physical Properties of Crystals*. Oxford University Press, Oxford, 1985.

[62] Tuckerman, M. E. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford University Press, Oxford, 2010.

[63] Cullity, B. D. and Stock, S. R. *Elements of X-Ray Diffraction*, 3rd ed. Prentice Hall, Upper Saddle River, 2001.