



Quantum-mechanical insights into sodium alginate gelation and AI-driven discovery of synthetic biomimetic analogs

K. P. Safna Hussan^{1,3} · S. Sooraj Kumar³ · Jibin K. Varghese⁴ · Rio Kita^{1,2} · Naoki Shinyashiki^{1,2}

Received: 14 April 2026 / Accepted: 26 August 2026
© The Author(s) 2026

Abstract

The functional versatility of sodium alginate (SA) in biomedical engineering is fundamentally governed by its sequence-dependent interaction with metal ions; however, the molecular mechanisms underlying these interactions remain incompletely understood. Herein, we establish a comprehensive molecular-level framework for SA and its synthetic biomimetic analogs by integrating density functional theory (DFT), artificial intelligence (AI)-assisted molecular design, molecular docking, molecular dynamics (MD) simulation, and MM-PBSA free energy analysis. Quantum chemical calculations revealed that β -D-mannuronic acid is slightly more stable than α -L-guluronic acid, while progressive polymerization and Na^+ coordination significantly enhanced structural stability and reduced the HOMO–LUMO gap from 10.75 to 10.96 eV in monomers to 8.15 eV in SA, accompanied by increased softness (0.123 eV^{-1}) and electrophilicity (2.031 eV). Theoretical FTIR spectra closely matched experimental observations, validating the optimized computational model. Guided by these insights, AI-assisted screening identified synthetic biomimetic scaffolds with substantially improved physicochemical and pharmacokinetic properties. Among the designed candidates, Analog 5 emerged as the closest structural mimic of SA, exhibiting reduced molecular weight ($234.20 \text{ g mol}^{-1}$), lower topological polar surface area ($102.29 \text{ \AA}^2$), zero Lipinski violations, and enhanced predicted bioavailability. Molecular docking against Aurora kinase A (AURKA) demonstrated the strongest binding for SA (Glide score = -6.142 ; IFD score = -11.23), followed by Analog 5 (-5.405 ; -7.97), Analog 4 (-4.496 ; -7.77), and Analog 3 (-4.147 ; -7.10). Subsequent 300 ns MD simulations confirmed stable ligand–protein complexes throughout the simulation, while MM-PBSA calculations revealed thermodynamically favorable binding free energies of $-666.260 \pm 1.291 \text{ kJ mol}^{-1}$ for SA, $-232.137 \pm 0.772 \text{ kJ mol}^{-1}$ for Analog 3, and $-47.735 \pm 0.315 \text{ kJ mol}^{-1}$ for Analog 5. Collectively, this integrated computational framework demonstrates that AI-designed biomimetic analogs successfully balance favorable molecular recognition with improved drug-like properties, providing a rational platform for developing next-generation synthetic alginate-inspired biomaterials for drug delivery, wound healing, tissue engineering, and regenerative medicine.

Keywords Sodium alginate · Density functional theory · Molecular docking · VEGF · AI-assisted molecular design · ADME prediction · Wound healing biomaterials

Introduction

Alginates are naturally occurring linear polysaccharides composed of β -D-mannuronic acid (M) and α -L-guluronic acid (G) units arranged in varying sequences along the polymer backbone. These units form homopolymeric blocks (GG and MM) and heteropolymeric sequences (MG/MG), and the relative abundance and arrangement of these blocks critically determine the physicochemical behavior of the polymer. Alginate is widely sourced from brown seaweeds and can also be produced by microbial fermentation, making it economically accessible in large quantities. Because of its

✉ K. P. Safna Hussan
yfm4476@tokai.ac.jp

¹ Micro/Nano Technology Center, Tokai University, Hiratsuka-shi, Kanagawa 259-1292, Japan

² Department of Physics, School of Science, Tokai University, Hiratsuka-shi, Kanagawa 259-1292, Japan

³ Ayurgreen Scientifica Research Institute, Ayurgreen Campus, Edapal, Malappuram, Kerala 679 582, India

⁴ Department of Chemistry, SH College, Thevara, Kochi, Kerala 682013, India

exceptional biodegradability, non-toxicity, tunable material properties, and biocompatibility, alginate is already used in a wide range of industrial applications, including advanced fields such as biomedicine and bioengineering [1]. In industrial and biomedical contexts, alginate is used as a thickener and stabilizer in the food industry, a carrier for controlled drug delivery, a scaffold material in tissue engineering, and as a matrix for cell encapsulation. The polymer's sol–gel transition under mild conditions is especially leveraged in applications such as 3D bioprinting, wound dressings, and agriculture [1] (Fig. 1).

A hallmark feature of alginate is its ability to form hydrogels in the presence of multivalent cations, particularly calcium (Ca^{2+}). This gelation is widely interpreted through the classical “egg-box” model, whereby Ca^{2+} ions selectively coordinate with G-block sequences, facilitating cross-linking of adjacent polymer chains and the formation of an extended three-dimensional network. The resulting gel's mechanical strength, porosity, and elasticity are highly sensitive to the M/G ratio and block architecture: alginates with a high proportion of G units tend to form stiffer and more brittle gels, whereas higher M content yields softer and more elastic materials [2].

Several studies have employed density functional theory (DFT) to elucidate alginate interactions with solvent molecules and metal cations at the molecular level. DFT calculations, including DFT-D3 dispersion corrections, have examined hydrogen bonding between alginate (and sodium alginate (SA) fragments) and increasing numbers of explicit water molecules, revealing progressively strengthened hydrogen-bond networks upon hydration. These studies show that SA interacts with water through a combination of medium-strength hydrogen bonds and metal–nonmetal interactions, with NBO and AIM/RDG analyses confirming the increasing contribution of hydrogen bonding and highlighting the necessity of explicit microsolvation to capture

realistic alginate–water interactions [3]. DFT has also been applied to alginate fragments complexed with calcium and other divalent cations, providing quantum-chemical support for features of the egg-box model. In particular, DFT studies on diuronate analogs complexed with Mg^{2+} , Ca^{2+} , Sr^{2+} , and transition metals demonstrate strong ionic and covalent coordination, with binding strength and geometry dependent on ion identity and hydration state, yielding detailed interaction energies, optimized structures, and orbital characteristics [4]. Furthermore, combined THz-TDS experiments and DFT calculations on SA under Ca^{2+} addition reveal that initial Ca^{2+} -mediated association preferentially involves intermolecular chain interactions, leading to crosslinked, near-tetrahedral coordination geometries around Ca^{2+} [5].

In recent years, integrated in-silico approaches have become indispensable for investigating biomolecular recognition and accelerating the rational design of therapeutic molecules and biomaterials. Techniques such as binding-site identification, molecular docking, molecular dynamics (MD) simulations, and binding free-energy calculations have been successfully applied to characterize transporter proteins, elucidate nanoparticle–protein interactions, and investigate enzyme inhibition mechanisms in drug discovery [6–8]. These computational strategies provide valuable molecular-level insights into binding mechanisms, structural stability, and interaction energetics while reducing the time and cost associated with experimental screening. Motivated by these advances, the present study integrates DFT calculations, AI-assisted molecular design, molecular docking, MD simulations, and MM-PBSA analyses to investigate SA and its synthetic biomimetic analogs [9]. Moreover, electronic descriptors are rarely linked quantitatively to macroscopic gelation behavior or synthetic mimetic design.

This computational investigation forms part of a broader study on amorphous solid dispersions (ASDs) of an anti-cancer drug, in which SA serves as the polymeric carrier and has been characterized experimentally by UV-visible, FTIR, broadband dielectric spectroscopy, and thermal analysis (TGA & DSC). SA is particularly suited to this role, as its water-soluble, un-cross-linked form maintains a homogeneous amorphous matrix and preserves free carboxylate groups for drug–polymer interaction features lost upon divalent cross-linking in calcium alginate. As a preliminary communication of the computational component, we address this critical knowledge gap by establishing a molecular-level framework that explicitly links alginate monomer identity, block sequence architecture, and cation coordination to gelation-relevant physicochemical properties. A hierarchical density functional theory (DFT) strategy was employed to systematically investigate monomeric units, representative block motifs, and extended oligomeric structures, enabling mechanistic insight into sequence-dependent



Fig. 1 Photographic image of alginates

electronic evolution. In parallel, to overcome the inherent limitations of natural alginate, including batch-to-batch variability and limited mechanical tunability, a biomedical AI-Q Research Agent was utilized to identify and evaluate synthetic biomimetic analogs. The shortlisted candidates were further assessed through pharmacokinetic (ADME) prediction and molecular docking, molecular dynamic simulations against Aurora kinase A to establish biological relevance and therapeutic applicability.

Computational methods

Modeling hierarchy

The computational modeling of SA was constructed using a hierarchical approach to systematically capture sequence- and length-dependent effects. At the most fundamental level, isolated mannuronate (M) and guluronate (G) monomers were modeled to establish intrinsic electronic structure, charge distribution, and coordination tendencies. The initial geometries of the M and G monomers were retrieved from the PubChem database and used as starting structures for full optimization. Building on this, dimeric units (MM, GG, and MG) were examined to probe nearest-neighbor effects, including sequence-dependent hydrogen bonding, conformational preferences, and initial cation-binding motifs. For each dimer (MM, GG, and MG), two distinct starting configurations were constructed and independently optimized, and the lower-energy conformer was retained as the representative structure.

Next, a representative block motif (MGMMGG) was constructed to capture short-sequence heterogeneity in natural alginates and to assess cooperative interactions within mixed blocks. Six configurations of this block were generated by sampling two orientations at each inter-residue junction, and the lowest-energy converged structure was selected, followed by ionization with sodium cations to generate SA. For the sodium-coordinated model, three configurations differing in Na⁺ placement relative to the carboxylate groups were optimized, and the most stable structure was retained. In all cases, vibrational frequency analysis confirmed the absence of imaginary frequencies, verifying that the selected geometries correspond to true energy minima. This block was then polymerized into an extended oligomer comprising six repeating SA units.

DFT functionals and basis sets

All quantum chemical calculations were carried out using DFT with the Gaussian suite of programs [10]. The structures of alginate monomers, dimers, and block models were

fully optimized employing the ω B97XD functional [11], which incorporates long-range corrections and empirical dispersion to account for hydrogen bonding and noncovalent interactions. The def2-GDZVP basis set [12] was used for all atoms. Vibrational frequency calculations were performed to verify that the optimized structures are true minima and to provide zero-point energies. All calculations were performed in the gas phase to establish the intrinsic electronic structure of the alginate framework as a baseline reference; explicit hydration and solvent effects will be addressed in a separate, dedicated study.

Fourier Transform Infrared (FTIR) spectroscopy analysis

FTIR spectra of SA were recorded at ambient temperature (296 K) using a JASCO FT/IR-4200ST spectrophotometer within the wavenumber range of 4000–500 cm⁻¹. For spectral acquisition, the samples were homogeneously mixed with spectroscopic-grade potassium bromide (KBr) and compressed into transparent pellets to ensure uniform sample distribution and enhanced infrared signal transmission. Each spectrum was obtained by averaging 16 consecutive scans at a spectral resolution of 4.0 cm⁻¹ to improve the signal-to-noise ratio. The experimental uncertainty associated with the recorded wavenumbers was estimated to be within ± 4.0 cm⁻¹.

Physicochemical property mapping

The NVIDIA MolMIM (Molecular Latent Space Informative Molecular Mapper) [13–15] architecture was employed to identify synthetic analogs of SA. The natural monomer was embedded into a 512-dimensional latent space. A Covariance Matrix Adaptation Evolution Strategy (CMA-ES) was utilized to navigate this space, optimizing for a target partition coefficient (plogP) of approximately -1.5 to -3.0, simulating the hydrophilic nature of polyuronic acids [16].

The synthetic accessibility (S_A) score of the selected analog was evaluated using the method of Ertl and Schuffenhauer [17], which combines fragment-based contributions with molecular complexity penalties to estimate the ease of chemical synthesis. The calculation was performed in RDKit using the SMILES representation of the optimized structure. Briefly, the synthetic accessibility score is derived from a fragment contribution term (S_{frag}) and a complexity penalty term (P_{complex}):

$$S_{\text{raw}} = S_{\text{frag}} - P_{\text{complex}}$$

where the fragment score is calculated from the occurrence frequencies of molecular fragments in large chemical databases:

$$S_{frag} = 1/N \sum_{i=1}^N C_i$$

Here, C_i represents the contribution of the i^{th} fragment and N is the total number of fragments identified in the molecule. The complexity penalty accounts for structural features that increase synthetic difficulty, including molecular size, stereocenters, ring complexity, bridgehead atoms, and macrocyclic structures:

$$P_{\text{complex}} = P_{\text{size}} + P_{\text{stereo}} + P_{\text{ring}} + P_{\text{bridge}} + P_{\text{macro}}$$

The resulting raw score is subsequently transformed and normalized to yield a final synthetic accessibility score ranging from 1 to 10:

$$S_A = f(S_{\text{raw}})$$

where lower values indicate compounds that are easier to synthesize and higher values correspond to structures with greater synthetic complexity. In general, compounds with SA scores below 3 are considered readily synthesizable, scores between 3 and 6 indicate moderate synthetic difficulty, and scores above 6 suggest challenging synthetic routes.

Physicochemical and pharmacokinetic property prediction

The physicochemical and pharmacokinetic properties of SA and the selected AI-derived analogs (Analog 5, Analog 4, and Analog 3) were predicted using the SwissADME online platform (<http://www.swissadme.ch>), which provides validated *in silico* models for drug-likeness and ADME profiling. Molecular structures were prepared in SMILES format and uploaded individually to the SwissADME server for analysis [18]. Key physicochemical descriptors, including molecular weight (MW), consensus octanol/water partition coefficient (LogP), topological polar surface area (TPSA), number of hydrogen bond donors (HBD) and acceptors (HBA), number of rotatable bonds, and Lipinski's rule-of-five compliance, were calculated to assess drug-likeness and oral bioavailability potential. Pharmacokinetic parameters were predicted using the platform's built-in models, including gastrointestinal (GI) absorption, blood–brain barrier (BBB) permeability, P-glycoprotein (P-gp) substrate probability, cytochrome P450 (CYP450) enzyme inhibition potential, bioavailability score, and skin permeability coefficient (log Kp). In addition, the brain or Intestinal

estimated permeation (BOILED-Egg) model was employed to graphically assess the relationship between lipophilicity (WLOGP) and polarity (TPSA), enabling rapid visualization of human intestinal absorption (HIA) and BBB penetration probability.

Target protein, ligand preparation and molecular docking

Molecular docking were performed using the Schrödinger Maestro molecular modeling suite (Release 2023-2, Schrödinger, LLC, New York, USA) to investigate the binding interactions of SA and the AI-designed analogs with human AURKA kinase domain (PDB ID: 8JF4).

Preparation of the target protein

The crystal structure of the human Aurora kinase A (AURKA) catalytic domain in complex with its co-crystallized ligand (ligand code C0N) was retrieved from the RCSB Protein Data Bank (PDB ID: 8JF4; resolved residues 127–389) [19]. As the deposited structure lacked electron density for residues 278–290, the complete kinase domain was reconstructed using AlphaFold3 [DOI: <https://doi.org/10.1038/s41586-024-07487-w>] to restore the missing loop region and generate a continuous model suitable for docking.

The stereochemical quality of the modeled structure was assessed using several independent validation tools, including VERIFY3D [20], ERRAT [21], SWISS-MODEL [22] QMEAN4, and the Ramachandran plot [23]. The stereochemical agreement between the modeled and experimental structures was further verified in Discovery Studio Visualizer [24]. The validated AlphaFold3 model was then subjected to a 50 ns molecular dynamics (MD) simulation to confirm its structural and energetic stability prior to docking.

Ligand preparation and molecular docking

Ligand geometries were first optimized in Gaussian 16 [10] and subsequently processed with the Schrödinger LigPrep module [25] to assign protonation states, generate low-energy conformers, and preserve stereochemistry. Docking was performed using the Schrödinger Glide module in Standard Precision (SP) mode [26, 27]. The docking grid was centered on the Cartesian coordinates of the co-crystallized ligand ($x = -7.69$, $y = -7.57$, $z = -6.18$). The resulting docked complexes were carried forward for further MD analysis.

Molecular dynamic simulation

Molecular dynamics (MD) simulations were performed using the best-scoring protein–ligand complexes obtained from the molecular docking studies. Simulations were carried out with GROMACS 2024.5 [28] using the OPLS [29] all-atom force field for the protein–ligand systems. Ligand topologies and partial atomic charges were generated using the LigParGen server [30–32], which provides OPLS-compatible parameters for small organic molecules. The optimized docking poses of all ligands were protonated at physiological pH (7.4) prior to topology generation.

Each protein–ligand complex was placed in a cubic simulation box, neutralized by adding appropriate Na^+ counterions, and solvated with the SPC/E water model. The solvated systems were subjected to energy minimization using the steepest-descent algorithm to remove unfavorable steric contacts [33]. Subsequently, two equilibration steps consisting of 2 ns NVT and 2 ns NPT simulations were performed with positional restraints applied to the heavy atoms of the protein–ligand complex, allowing the solvent and ions to equilibrate around the restrained complex. These equilibration steps ensured stable temperature, pressure, and density before the production simulation.

The equilibrated systems were then subjected to 300 ns production MD simulations with a 2 fs integration time step under periodic boundary conditions. The temperature (300 K) and pressure (1 bar) were maintained using the Berendsen thermostat and Berendsen barostat [34], respectively. Short-range non-bonded interactions were calculated using the Verlet neighbor-list scheme, while long-range electrostatic interactions were treated using the Particle Mesh Ewald (PME) method [35].

The structural stability and conformational dynamics of the protein–ligand complexes were evaluated by analyzing the root mean square deviation (RMSD), root mean square fluctuation (RMSF), and radius of gyration (R_g) throughout the simulation. Hydrogen-bonding interactions and other non-covalent contacts during the simulation trajectory were analyzed using the built-in GROMACS analysis utilities together with the Protein–Ligand Interaction Profiler (PLIP) tool [36]. To further quantify the binding affinity, MM-PBSA binding free energy calculations were performed using the g_{mmpbsa} package. A total of 1000 snapshots extracted from the equilibrated 280–290 ns segment of the trajectory at 10 ps intervals were used to calculate the van der Waals, electrostatic, polar solvation, solvent-accessible surface area, and total binding free energy components.

MM-PBSA binding free energy calculations

The Molecular Mechanics Poisson–Boltzmann Surface Area (MM-PBSA) method, originally developed by Kollman and co-workers, is a widely used approach for estimating the binding free energy of protein–ligand complexes from molecular dynamics trajectories [37]. Owing to its balance between computational efficiency and accuracy, MM-PBSA has been extensively applied in drug discovery, binding affinity prediction, conformational stability analysis, and identification of key residues contributing to molecular recognition.

In the present study, binding free energy calculations were performed using the g_{mmpbsa} package [38], which is compatible with GROMACS molecular dynamics trajectories. The energy components corresponding to the protein (P), ligand (L), and protein–ligand complex (C) were extracted using the single-trajectory approach, thereby minimizing conformational sampling errors. Binding free energies were calculated from 1000 snapshots extracted from the equilibrated region of the 300 ns MD trajectory (280–290 ns) at 10 ps intervals. The total binding free energy (ΔG_{bind}) in the solution phase was computed as the difference between the free energy of the protein–ligand complex and the sum of the free energies of the isolated protein and ligand, as expressed in Eq. (1). The total binding energy was further decomposed into van der Waals (vdW), electrostatic (ESE), polar solvation (PSE), and nonpolar solvent-accessible surface area (SASA) energy contributions to provide detailed insights into the molecular forces governing ligand binding.

$$\Delta G_{\text{bind}} = G_c - G_p + G_L$$

Results and discussion

Quantum chemical optimization and energetic stability

Quantum chemical optimization and total energy calculations were conducted on alginate monomers, oligomeric blocks, and SA models to assess relative stability and structural changes with polymerization and ionic substitution. Optimized geometries and total energies (E , kJ/mol) are presented in Fig. 2.

Among monomers, β -D-mannuronic acid (M) was slightly more stable ($E = -1,998,257.29$ kJ/mol) than α -L-guluronic acid (G; $E = -1,998,251.23$ kJ/mol), attributable to stereochemical differences at C-5 affecting hydroxyl orientation and intramolecular hydrogen bonding. This aligns with greater chain flexibility in M-rich

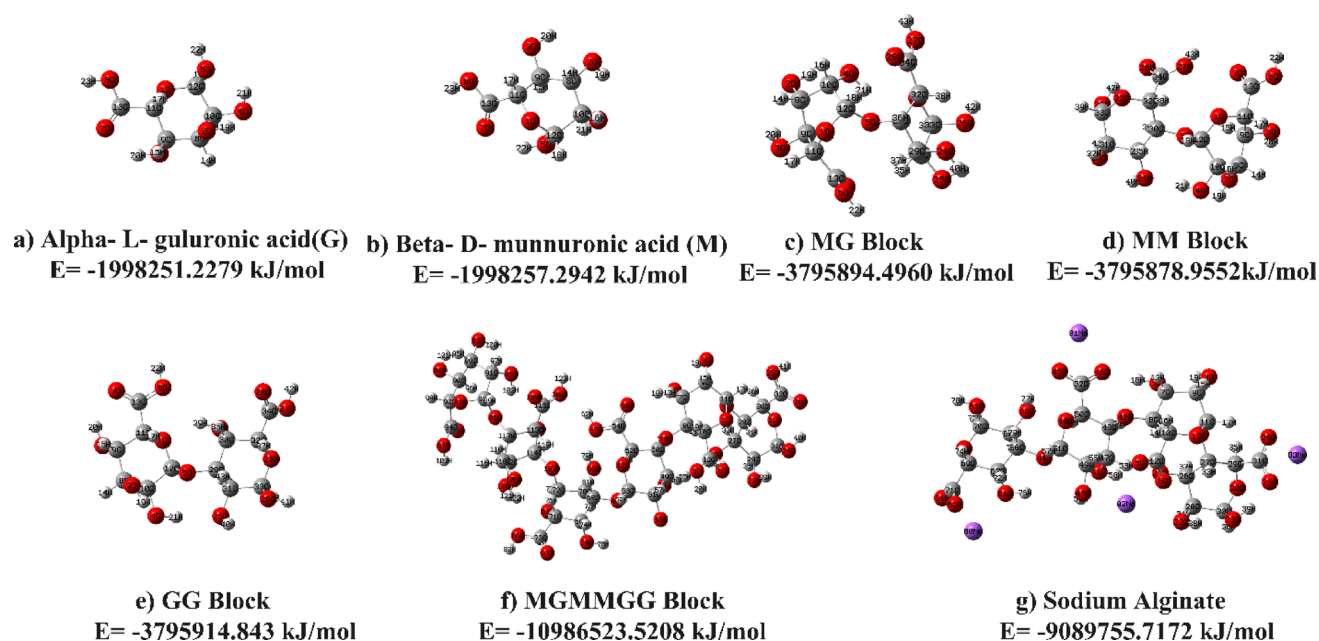


Fig. 2 Optimized molecular structures and relative energy profiles of alginate monomeric units, oligomeric blocks, and SA polymer. **a)** α -L-guluronic acid (G), **b)** β -D-mannuronic acid (M), **c)** MG block, **d)**

MM block, **e)** GG block, **f)** MGMMGG block, and **g)** SA polymer are shown after geometry optimization

Table 1 Calculated binding energies (ΔE_{bind}) of the sodium ion (Na^+) with the GG, MG, and MM dimer blocks at the ω B97XD/def2-TZVP level

Complex	$E_{(\text{complex})}$ kJ/mol	$E_{(\text{block})}^-$ kJ/mol	$E_{(\text{Na}^+ \text{ ion})}$ kJ/mol	ΔE_{bind} kJ/mol
NaGG	-4220523.799515	-3794526.916635	-425467.700187	-529.182693
NaMG	-4220554.764664	-3794508.131181	-425467.700187	-578.933296
NaMM	-4220577.472616	-3794585.058337	-425467.700187	-524.714092

regions. Among the dimers, the GG block was the most stable ($E = -3,795,914.85$ kJ/mol), followed by MG ($-3,795,894.49$ kJ/mol) and MM ($-3,795,878.96$ kJ/mol). The enhanced stability of the GG sequence is consistent with the cooperative hydrogen-bonding and electron-dense carboxylate environment identified in the frontier orbital and electrostatic potential analyses (Sects. 3.2–3.4), which also underlies its established selectivity for divalent cation coordination. To directly assess the cation-coordination preference underlying gelation, binding energies were computed for Na^+ with each dimer block (GG, MG, MM) as $\Delta E_{\text{bind}} = E_{(\text{complex})} - [E_{(\text{block})}^- + E_{(\text{Na}^+ \text{ ion})}]$, with counterpoise correction for basis-set superposition error. The resulting values (Table 1) are -529.18 , -578.93 , and -524.71 kJ/mol for the NaGG, NaMG, and NaMM complexes, respectively, indicating strong and energetically favorable Na^+ coordination in all three cases. MG binds Na^+ most strongly (-578.93 kJ/mol), then GG (-529.18 kJ/mol), then MM (-524.71 kJ/mol). Notably, this ordering differs from the intrinsic stability ordering of the bare blocks ($\text{GG} > \text{MG} > \text{MM}$), indicating that monovalent Na^+ coordination is not governed by block stability alone but reflects the local carboxylate geometry and charge distribution available to the cation.

Longer chains exhibited substantial stabilization: the MGMMGG hexamer had $E = -10,986,523.52$ kJ/mol, confirming enhanced stability with polymerization degree, which supports alginate's gel-forming properties. The SA model ($E = -9,089,755.72$ kJ/mol) revealed significant stabilization via Na^+ coordination to carboxylate groups, reducing electrostatic repulsion and modulating chain conformation—consistent with experimental dependence on counterions and salt concentration.

These findings establish a structure–energy relationship governed by stereochemistry, sequence, chain length, and ionic interactions, providing molecular insights into alginate's physicochemical behavior and aiding rational design for biomedical applications such as drug delivery and wound dressings.

Frontier molecular orbital (FMO) analysis

The frontier molecular orbitals — the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) — govern the electron-donating and electron-accepting capacity of a molecule, and the energy separation between them (the HOMO–LUMO gap,

$\Delta E = E_{LUMO} - E_{HOMO}$) serves as a key indicator of kinetic stability, chemical reactivity, and charge-transfer propensity. A larger gap corresponds to greater stability and lower reactivity, whereas a smaller gap denotes enhanced polarizability and a stronger tendency toward charge redistribution and intermolecular interaction. This descriptor has been widely applied to characterize the reactivity and binding behavior of π -conjugated and macrocyclic surfaces in recent computational studies [39–41] (Fig. 3).

Frontier molecular orbital (FMO) analysis was performed to understand the electronic structure evolution of alginate from individual monomeric units to extended block sequences and finally to SA. For the G unit, the HOMO is primarily localized on the ring and carboxyl carbon atoms, and the LUMO is distributed over the oxygen-rich carboxylate and hydroxyl groups, suggesting potential electrophilic interaction sites. In the M unit (mannuronic acid), the HOMO shows greater delocalization along the hydroxyl group of the ring the LUMO remains distributed over the carboxyl group of the ring system. For the MG block, the alternating sequence leads to noticeable delocalization of both HOMO and LUMO across adjacent M and G residues. This electronic communication between the two units results in a modified electronic environment with intermediate reactivity. The MM block exhibits stronger HOMO delocalization along the mannuronic backbone, reflecting increased electron-donating character. In contrast, the GG block shows HOMO density concentrated around closely spaced carboxylate oxygen atoms, forming electron-rich pockets favorable for cation binding. The LUMO distribution supports effective charge acceptance in these regions. With increasing sequence length, as observed in the MGM-MMG block, HOMO is located over electron-rich GG block and LUMO become extensively delocalized over the MG block of the oligomer.

A systematic narrowing of the frontier orbital energy gap ($\Delta E = E_{LUMO} - E_{HOMO}$) is evident in the models of alginate systems as structural complexity increases from isolated monomeric units to oligomeric blocks and ultimately to the polymeric SA form. The G unit exhibits a ΔE of 10.75 eV, M unit shows a slightly higher value of 10.96 eV, reflecting highly localized FMO characteristic of small, insulated molecular systems with limited π -conjugation and strong electron confinement. In dimeric blocks, orbital delocalization begins to emerge, leading to modest gap reductions. The heteromeric MG block yields ΔE =10.95 eV, and the homomeric MM block 10.84 eV, attributable to extended electronic coupling along the mannuronate backbone. A more substantial decrease occurs in the GG block (ΔE =10.27 eV), consistent with enhanced cooperative effects and hydrogen-bonding networks between adjacent guluronate residues, which facilitate greater orbital overlap

and align with the established selectivity of GG sequences for divalent cation coordination. Extension to longer oligomers, such as the MGMMGG hexamer, maintains $\Delta E \approx 10.27$ eV, suggesting saturation of collective electronic effects in moderate chain lengths and the emergence of quasi-polymeric behavior. The most significant reduction arises in the fully polymeric SA structure, where carboxylate deprotonation and Na^+ counterion coordination yield ΔE =8.15 eV. This transition from neutral, discrete oligomers to an ionic polyelectrolyte introduces significant electrostatic stabilization of the LUMO, enhanced polarizability, and charge delocalization across the extended chain.

Global reactivity descriptors

The progressive narrowing of the frontier orbital gap correlates with increased chemical softness, polarizability, and ion-binding affinity, accounting for the superior functional properties of SA over its oligomeric precursors. This is driven by polymerization and ionization, as quantified by global reactivity descriptors from Koopmans' theorem (Table 2) [42, 43].

The ionization energy (IE) values for the G (9.425 eV) and M (9.658 eV) units are relatively high, indicating strong resistance to electron removal and electronic stability at the monomer level. Correspondingly, their electron affinity (EA) values are moderately negative (− 1.330 and − 1.338 eV), suggesting limited electron-accepting ability. Upon block formation, subtle but systematic changes occur: the MG, MM, and GG blocks show slightly reduced IE values (9.444–9.635 eV) and marginal variations in EA, reflecting increased orbital delocalization and enhanced electronic communication between neighboring residues. The extended MGMMGG block exhibits the lowest IE among the neutral oligomers (9.317 eV), indicating facilitated electron donation as chain length increases. The electronegativity (χ) and chemical potential (μ) values further support this trend. While χ remains within a narrow range (~ 3.97 – 4.22 eV) for monomers and short blocks, the MGMMGG block shows a relatively higher χ (4.180 eV), suggesting increased ability to attract electrons due to cooperative interactions along the chain. The negative chemical potential values confirm the thermodynamic stability of all systems, with extended blocks exhibiting enhanced electronic softness.

Notably, the chemical hardness (η) decreases progressively from monomeric units (5.377–5.498 eV) to block structures, reaching 5.138 eV for the MGMMGG block, while the corresponding softness (S) increases (0.091–0.097 eV^{−1}). This trend indicates a gradual transition from hard, less reactive molecular systems to softer, more polarizable oligomers. The electrophilicity index (ω) also increases with chain growth, from 1.523 eV (G unit) and

Fig. 3 Frontier molecular orbital (FMO) distributions and HOMO–LUMO energy gaps of alginate monomeric units, block structures, and SA polymer. The spatial electron density distributions of the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) are shown for **a)** G unit, **b)** M unit, **c)** MG block, **d)** MM block, **e)** GG block, **f)** MGMMGG block, and **g)** SA

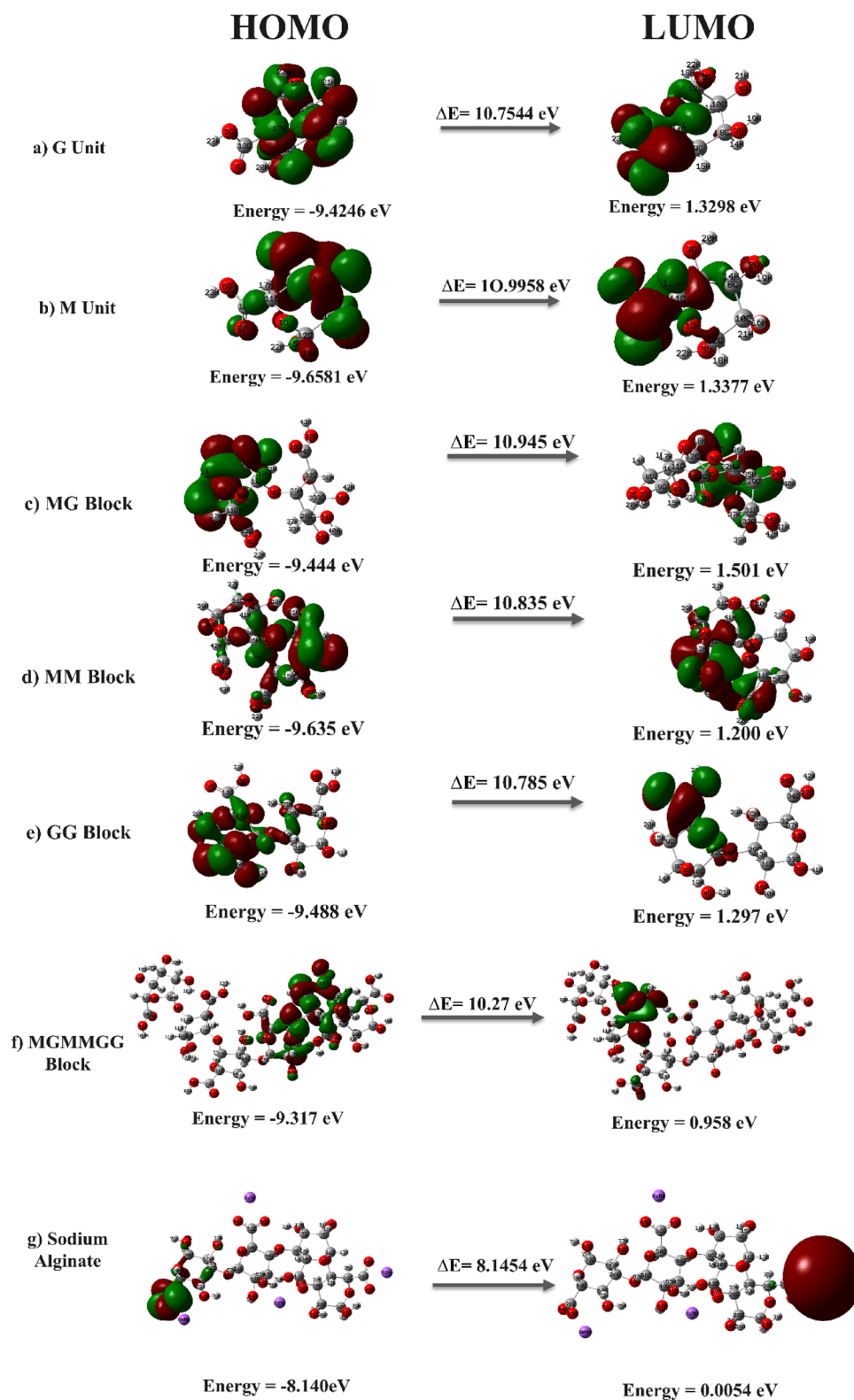


Table 2 Global reactivity descriptors illustrating the electronic evolution of alginate from monomeric units to SA

Sample	IE (eV)	EA (eV)	χ (eV)	μ (eV)	ω (eV)	η (eV)	S
a) G Unit	9.425	− 1.330	4.047	− 4.047	1.523	5.377	0.093
b) M Unit	9.658	− 1.338	4.160	− 4.160	1.574	5.498	0.091
c) MG Block	9.444	− 1.501	3.972	− 3.972	1.441	5.473	0.091
d) MM Block	9.635	− 1.200	4.218	− 4.218	1.642	5.418	0.092
e) GG Block	9.488	− 1.297	4.096	− 4.096	1.555	5.393	0.093
f) MGMMGG Block	9.317	− 0.958	4.180	− 4.180	1.700	5.138	0.097
g) SA	8.140	− 0.005	4.067	− 4.067	2.031	4.073	0.123

1.574 eV (M unit) to 1.700 eV for the MGMMGG block, highlighting enhanced ability to participate in charge-transfer interactions.

The most dramatic electronic transformation occurs upon formation of SA. The IE drops significantly to 8.140 eV, and the EA approaches zero (−0.005 eV), reflecting strong stabilization of added electrons due to ionic character. This is accompanied by a marked reduction in hardness ($\eta=4.073$ eV) and a pronounced increase in softness ($S=0.123$ eV^{−1}), along with the highest electrophilicity index ($\omega=2.031$ eV). These changes signify a transition from neutral covalent systems to a highly polar, ionic polymer with enhanced reactivity, ion-exchange capability, and interaction potential. Overall, the progressive decrease in hardness and increase in softness and electrophilicity from monomers to SA rationalize the superior functional performance of SA in gelation, metal-ion binding, and pharmaceutical applications.

Molecular electrostatic potential (MEP) mapping

The molecular electrostatic potential (MEP) surfaces clearly illustrate how the electronic environment of alginate evolves from isolated monomeric units to the ionic polymer, SA. In the G and M units, the MEP maps show well-defined, localized negative potential regions (red/yellow) concentrated around the carboxyl and hydroxyl oxygen atoms, while positive potential regions (blue) are mainly associated with hydrogen atoms [44]. Upon formation of MG, MM, and GG blocks, the negative potential regions become more extended and interconnected along the polymer backbone, indicating increased charge delocalization and cooperative electronic interactions between adjacent uronic acid residues. Among these, the GG block exhibits the most intense and contiguous negative potential zones due to the close spatial arrangement of carboxylate groups, highlighting its strong affinity for metal-ion coordination and supporting its dominant role in alginate gelation. As the chain length increases further in the MGMMGG block, the MEP surface shows pronounced spreading of negative electrostatic potential over the entire GG unit. The most significant transformation occurs upon conversion to SA. Deprotonation of carboxylic acid groups and coordination with Na⁺ ions dramatically reorganize the

electrostatic landscape. The negative potential becomes more uniformly distributed along the backbone, while localized positive regions emerge around the sodium ions, confirming the formation of strong ionic interactions. This redistribution reduces localized charge extremes, stabilizes the system electrostatically, and enhances overall polarity. Consequently, SA exhibits increased ionic character, higher softness, and superior ion-exchange capability compared to its molecular precursors, providing a clear electronic rationale for its enhanced solubility, gel-forming ability, and functional performance in biomedical and pharmaceutical applications (Fig. 4).

With increasing sequence length, as observed in the MGMMGG block, both HOMO and LUMO become extensively delocalized over the entire oligomer. The HOMO–LUMO gap further decreases, indicating progressive band-gap narrowing with chain growth. This trend reflects the transition from localized molecular behavior to collective electronic properties characteristic of polymeric systems.

Upon formation of SA, significant electronic restructuring occurs due to deprotonation of carboxylic acid groups and coordination with Na⁺ ions. The HOMO energy becomes less negative, indicating enhanced electron-donating capability, while the LUMO energy decreases dramatically toward zero. Consequently, the HOMO–LUMO gap is substantially reduced (~8.15 eV), signifying increased ionic character, polarity, and interaction potential. This pronounced gap reduction explains the improved solubility, ion-exchange capacity, and functional performance of SA in biological and pharmaceutical applications.

Overall, the sequential progression from monomeric units to block structures and finally to SA reveals a clear trend of increasing orbital delocalization and decreasing HOMO–LUMO gap. These electronic changes underpin the unique physicochemical properties of alginate, particularly its metal-ion binding, gelation behavior, and versatility as a biopolymeric material.

FTIR spectra

FTIR spectroscopy was employed to elucidate the progressive structural evolution of SA from its fundamental

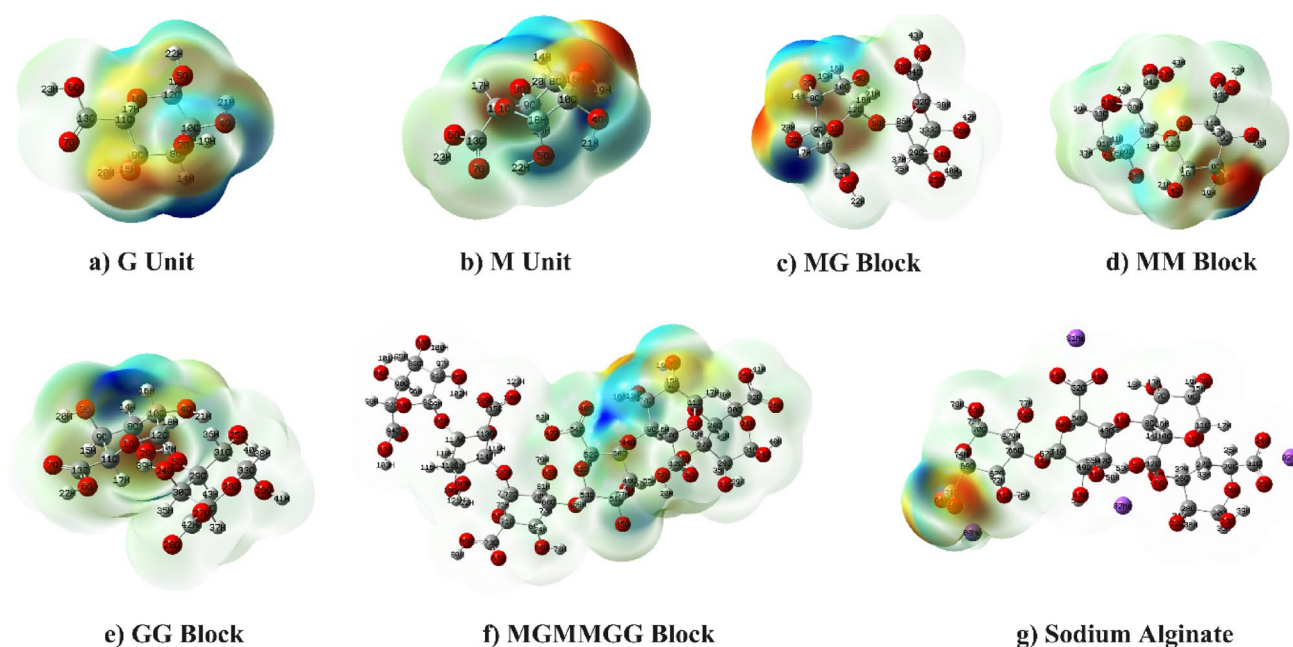
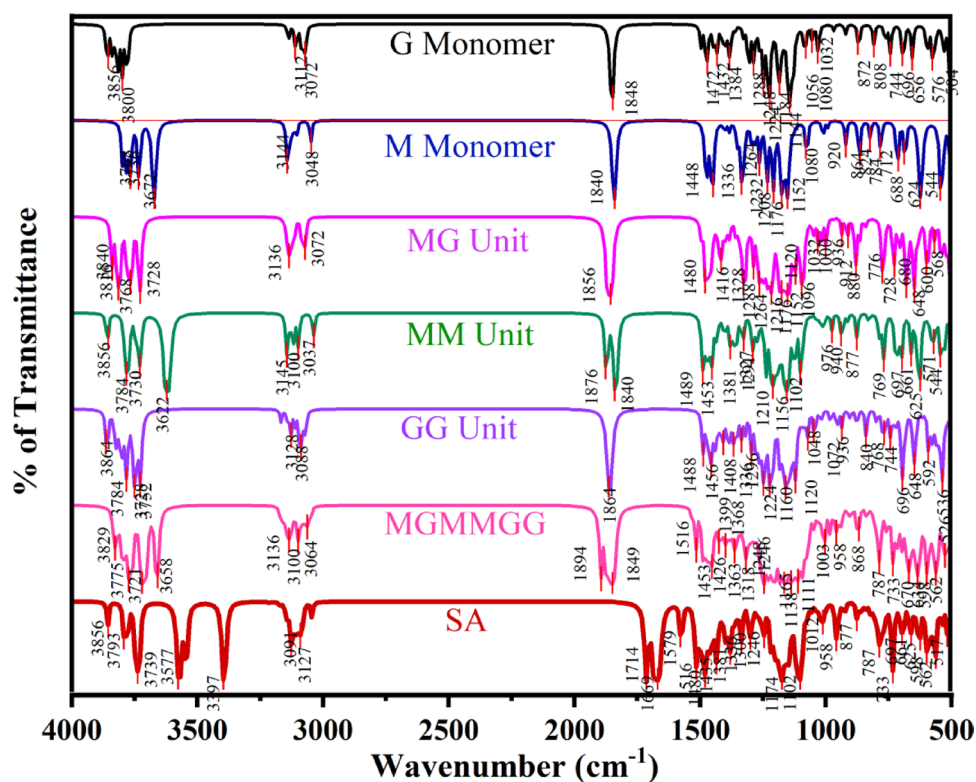


Fig. 4 Molecular electrostatic potential (MEP) surface maps of alginate monomeric units, block structures, and SA polymer. The MEP distributions are shown for **a)** G unit, **b)** M unit, **c)** MG block, **d)** MM block, **e)** GG block, **f)** MGMMGG block, and **g)** SA

Fig. 5 Theoretical FTIR Spectra of alginate monomeric units, block structures, and SA polymer



building blocks, namely guluronic (G) and mannuronic (M) acid units and their intermediate oligomeric assemblies (MG, MM, GG, and MGMMGG). The comparative spectra are presented in Fig. 5. All units show multiple sharp peaks (monomers) that coalesce into a broad, intense band in SA centered around 3397 cm^{-1} , indicating extensive

intermolecular hydrogen bonding in the polymer. Distinct peaks in G and M (e.g., $3072, 3144\text{ cm}^{-1}$) are conserved but slightly shifted in the MG/MM/GG units and finally appear as a doublet at $3127/3091\text{ cm}^{-1}$ in SA. Monomers show sharp peaks near $1840\text{--}1850\text{ cm}^{-1}$. In SA, a massive shift occurs to 1609 cm^{-1} and 1714 cm^{-1} , corresponding to

Table 3 Functional group assignments of each characteristic peak observed in FTIR spectra of alginate monomeric units, block structures, and SA polymer

Wave-number range (cm ⁻¹)	Observed in (Fig. 5)	Functional group assignment	Structural interpretation
3600–3200	G, M, MG, MM, GG, MGM-MGG, SA	Broad – OH stretching (hydrogen bonded hydroxyls)	Indicates polysaccharide backbone formation and extensive intermolecular hydrogen bonding in alginate polymer network
2920–2850	All samples	C–H stretching (– CH, – CH ₂)	Confirms carbohydrate ring framework remains conserved during polymer assembly
1740–1710	Weak/absent in SA	C=O stretching (COOH, protonated acid)	Reduction in free acid groups and conversion into sodium carboxylate salt in SA
1620–1590	Strong in SA and oligomers	Asymmetric stretching of – COO ⁻	Signature peak of alginate salt formation, confirms ionized carboxylate groups
1440–1410	Prominent in SA	Symmetric stretching of – COO ⁻	Confirms SA structure and polymer ionic stabilization
1320–1250	MG, MM, GG, SA	C–O stretching of carboxyl and pyranose ring	Demonstrates glycosidic polymer linkage evolution
1150–1080	All units and SA	C–O–C stretching (glycosidic bond)	Confirms polymerization between M and G monomers
1030–1020	SA and oligomers	C–O stretching (pyranose ring vibration)	Indicates polysaccharide ring stabilization
950–890	GG dominant	Guluronic acid fingerprint region	Stronger in GG-rich fractions → reflects block structure formation
890–820	MM dominant	Mannuronic acid fingerprint region	Indicates M-block contribution in copolymer
800–600	SA	Skeletal ring vibrations	Confirms long-chain polysaccharide backbone formation

asymmetric stretching of the carboxylate salt cm⁻¹. The fingerprint region shows the formation of the glycosidic linkage. SA displays a prominent peak at 1033 cm⁻¹, a merger of the monomeric ether/alcohol signals. M-units show a peak near 808–812 cm⁻¹, while G-units show peaks near 744 cm⁻¹. SA retains these as smaller markers at 877, 787 cm⁻¹ indicating the uronic acid configuration.

To validate the theoretical structure of SA, the theoretical vibrational modes (calculated or standard assignments) using DFT/ ω B97XD functional were compared with experimental FTIR spectra and shown in Fig. 6. The comparative analysis between simulated and experimental FTIR spectra

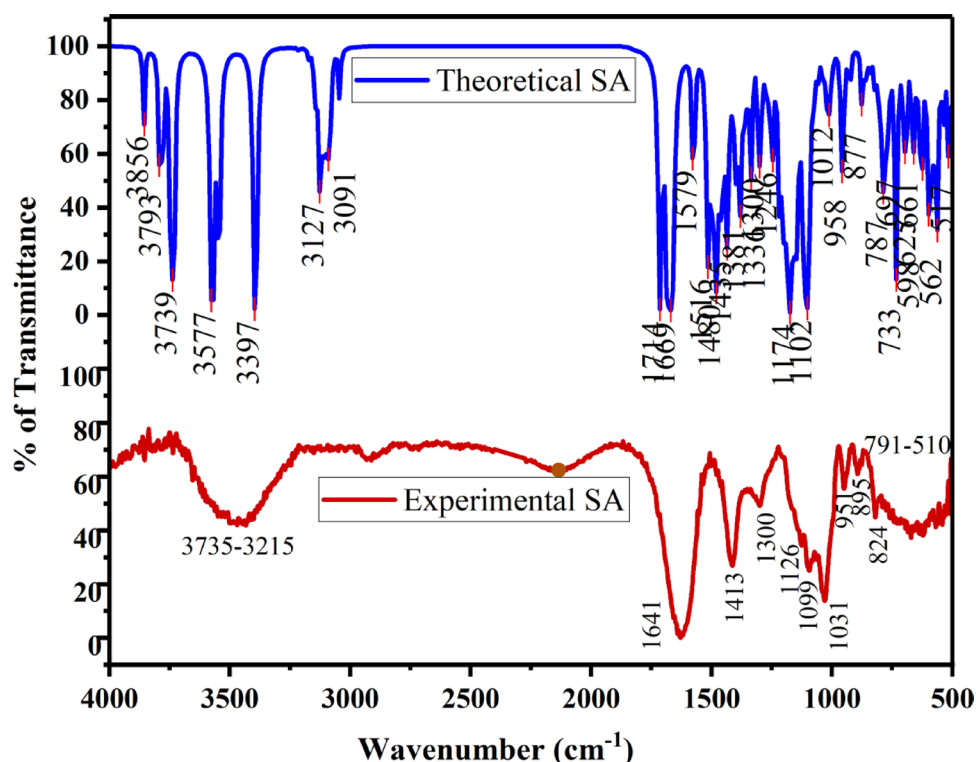
demonstrates strong agreement in both peak positions and functional group assignments, validating the accuracy of the optimized SA molecular model. The broad O–H stretching region shows a minor red shift in experimental spectra, which is attributed to extensive intermolecular hydrogen bonding and moisture interactions in the solid polymer state that are absent in isolated theoretical calculations. The characteristic asymmetric and symmetric stretching vibrations of carboxylate groups (–COO⁻) exhibit excellent correspondence between theoretical (1646 and 1461 cm⁻¹) and experimental (1641 and 1413 cm⁻¹) values, confirming the correct representation of ionic SA structure in the computational model. Similarly, the glycosidic linkage vibrations and pyranose ring modes appear consistently in both spectra, demonstrating preservation of the polysaccharide backbone architecture.

Minor frequency deviations observed in the fingerprint region can be attributed to solid-state packing effects, polymer chain entanglement, hydration level variations, and anharmonic vibrational contributions not fully captured in gas-phase theoretical simulations. Despite these expected differences, the close spectral overlap strongly supports the reliability of the computational structure used for subsequent electronic structure analysis and molecular property evaluation.

AI-driven biomimetic discovery

AI systematic evaluation was performed to identify synthetic analogs to the natural biopolymer SA, with a particular focus on structural similarity, physicochemical properties, and functional group distribution relevant to gelation behavior. The AI-assisted systematic evaluation identified Analog 5 as the superior biomimetic candidate for replicating the functionality of β -D-mannuronic and α -L-guluronic acid units. As shown in Table 2, Analog 5 exhibits a highly negative pLogP (– 3.15), indicating strong hydrophilicity comparable to natural SA. The synthetic feasibility of Analog 5 was further confirmed by its computed SA score of 3.49, which falls well within the readily accessible range. This low value is consistent with its structurally simple framework — a single tetrahydropyran ring bearing two carboxylic acid groups — and indicates that the candidate is synthetically tractable, although experimental synthesis and characterization remain necessary to confirm its accessibility and gelation behavior. Crucially, its tetrahydropyran core and di-carboxylic acid functionality provide the necessary spatial orientation for divalent cation chelation, mirroring the classical ‘egg-box’ mechanism. In contrast, Analogs 1 and 2 showed increasing lipophilicity (pLogP > – 0.5), rendering them unsuitable for hydrogel-based biomedical applications.

Fig. 6 Experimental and theoretical FTIR spectral comparison of SA for structural validation



Analog 5 exhibits a high degree of structural homology with the β -D-mannuronic acid repeating units of SA, primarily due to its tetrahydropyran backbone that closely reproduces the pyranose ring conformation of native alginate monomers. Preservation of this cyclic scaffold is critical for maintaining steric complementarity and conformational rigidity that govern supramolecular organization in polysaccharide-derived hydrogels. The incorporation of two strategically positioned carboxylic acid functionalities further enables multidentate coordination with divalent cations, facilitating calcium-mediated ionic crosslinking through an “egg-box”-type interaction mechanism. Furthermore, the elevated carboxylate group density in Analogs 4 and 5 is expected to increase metal ion binding affinity, providing a carboxylate arrangement geometrically compatible with egg-box-type divalent coordination at the monomer level, a prerequisite for — but not by itself sufficient for — network formation upon polymerization and crosslinking. The cyclic frameworks observed in Analogs 3 and 4 highlight the importance of ring-based scaffolds in emulating the pyranose units of alginate. Such rigid or semi-rigid backbones can enhance structural stability and resistance to deformation. Compared to native alginate, which is known to exhibit limited mechanical strength, these synthetic analogs—particularly those with higher molecular weights—may offer improved toughness and durability while retaining the essential gelling functionalities. Although Analog 3 demonstrates comparatively lower ionic crosslinking potential

than Analog 5, its conformational rigidity and balanced physicochemical profile suggest complementary reinforcement behavior when incorporated into composite hydrogel systems. Collectively, these structure–property relationships highlight Analog 5 as the primary gelling scaffold and Analog 3 as a mechanically stabilizing component, offering a synergistic design strategy for engineering next-generation alginate-inspired synthetic hydrogels with tunable mechanical and physicochemical performance.

AI-identified synthetic analogs represent a strategic shift from natural SA toward precision-engineered biomimetics. These synthetic systems outperform seaweed-derived materials by eliminating batch variability and allowing for programmable degradation and enhanced mechanical strength. Furthermore, synthetic analogs offer significant economic benefits; they are patentable as proprietary intellectual property and can be produced at scale from low-cost feedstocks, making them a more consistent and cost-effective alternative to traditional extraction methods. To bridge the structural advantages of AI-engineered materials with their functional performance, we evaluated the pharmacological potential of these synthetic candidates.

Physicochemical and pharmacokinetic properties prediction

The predicted physicochemical and pharmacokinetic properties of the synthesized compounds were evaluated using

Table 4 Comparison of theoretical and experimental FTIR vibrational bands of SA

Theoretical FTIR (cm ⁻¹)	Experimental FTIR (cm ⁻¹)	Band assignment	Correlation
3793–3127	3735–3215	O–H stretching (hydrogen bonded hydroxyl groups)	Excellent agreement; slight red shift due to strong intermolecular hydrogen bonding in solid sample
1646	1641	Asymmetric stretching of –COO ⁻	Strong correlation; confirms sodium carboxylate formation
1579	1513	Carboxylate coupled vibration	Moderate shift due to ion coordination environment
1461	1413	Symmetric stretching of –COO ⁻	Good agreement; confirms ionic alginate backbone
1390	1300	C–O stretching (carboxyl and pyranose ring)	Shift attributed to polymer packing effects
1330	1126	C–O–C glycosidic stretching	Slight deviation due to conformational flexibility
1174	1090	Glycosidic linkage vibration	Strong overlap indicating backbone consistency
1102	1031	Pyranose ring C–O stretching	Excellent agreement
1012	958	Ring breathing vibration	Good match
871	824	M/G fingerprint region	Consistent block structural fingerprint
733	791–510	Skeletal vibration modes	Broadened experimentally due to polymeric disorder

the SwissADME platform to assess their potential as oral drug candidates. Figure 7 illustrates the chemical structures and corresponding bioavailability radar profiles of SA and its selected analogs (Analog 5, Analog 4, and Analog 3). The radar plots summarize key physicochemical parameters, including lipophilicity, molecular size, polarity, solubility, flexibility, and degree of saturation, which are critical for evaluating drug-likeness. SA exhibits high polarity and solubility but comparatively limited lipophilicity, which may affect membrane permeability. In contrast, the analogs demonstrate a more balanced distribution of physicochemical properties, with improved lipophilicity and optimized molecular flexibility while maintaining acceptable polarity. Among the tested derivatives, Analog 3 and Analog 4 show enhanced overall profile uniformity, suggesting better pharmacokinetic suitability and potential oral bioavailability.

As shown in Table 6, the parent molecule SA exhibited two violations, primarily driven by an excessively high Topological Polar Surface Area (TPSA) of 215.50 Å² and an overabundance of hydrogen bond acceptors (13), suggesting suboptimal oral bioavailability and restricted pharmacokinetic adaptability. In contrast, Analogs 5, 4, and 3 show zero violations, with significant reduction in molecular weights (234–310 g/mol) and balanced LogP values (–0.50 to 1.77) while SA has –4.18, indicating significantly higher potential for oral bioavailability. The dramatic reduction in TPSA (from 215.50 Å² in SA to as low as 64.99 Å² in Analog 3) further supports the likelihood of passive intestinal absorption, as TPSA values below 140 Å² (and ideally below 90 Å²) are strongly correlated with better cell membrane penetration.

Table 5 AI-assisted systematic evaluation and biomimetic scoring of synthetic analogs compared to natural SA

Index	Type	SMILES structure	Solubility (plogP)	Chemical functionality	Alginate similarity
SA	Natural	OC1C(O)C(OC(O1)C(=O)O)OC2C(O)C(O)C(OC2C([O-])=O)O	~ – 3.00	Polysaccharide / Carboxyl	Baseline (100%)
5	Synthetic	CC(OC[C@@H]1COCCO1)(C(=O)O)C(=O)O	– 3.15		
4	Synthetic	COC1(COC(=O)[C@@H]2 C[C@@H]2 C(=O)O)CCOCC1	– 2.0239	Cyclic Ether / Acid	High match: Similar “ring” geometry to sugar monomers.
3	Synthetic	O=C(O)C1(C(=O)OC2(C3CCOCC3)CCOCC2)COC1	– 1.5002	Complex Ester / Acid	Moderate: Good solubility but complex backbone.
2	Synthetic	COc1ccc(C(=O)O)c(OC(=O)[C@H]2 C(C)(C)O[C@H]2 C)c1	– 0.4671	Aromatic / Acetal	Low: Approaching “fat-soluble” territory.
1	Synthetic	COc1ccc(OC2(C(=O)O)CCOCC2)cc1	0.5155	Aromatic Ether	Very low: Lipophilic; unlikely to form hydrogels.

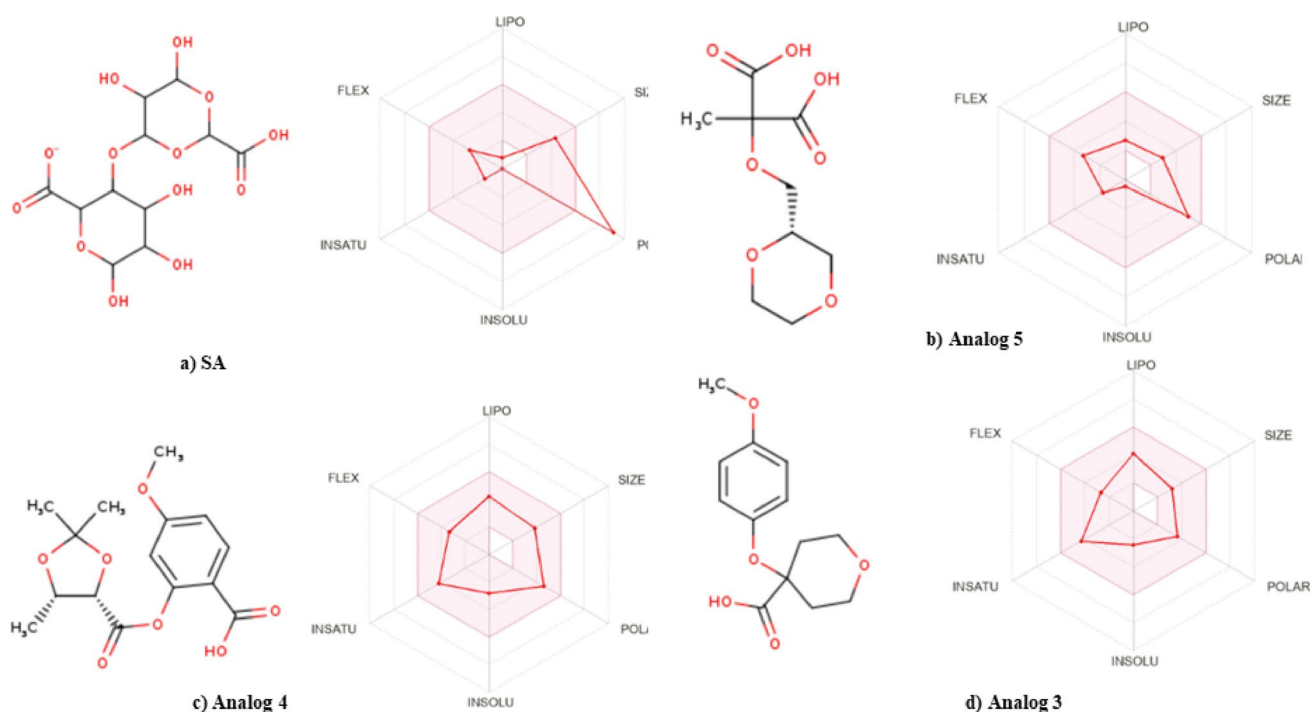


Fig. 7 Chemical structures and bioavailability radar plots of **a)** SA, **b)** Analog 5, **c)** Analog 4, and **d)** Analog 3

Table 6 Physicochemical properties and drug-likeness

Molecule	MW (g/mol)	Log Po/w (Consensus)	TPSA ($\text{\AA}^2$)	H-bond donors	H-bond acceptors	Rotatable bonds	Lipinski violations
SA	355.23	-4.18	215.50	6	13	4	2
Analog 5	234.20	-0.50	102.29	2	7	5	0
Analog 4	310.30	1.77	91.29	1	7	5	0
Analog 3	252.26	1.61	64.99	1	5	4	0

Table 7 Pharmacokinetic and ADME profile

Molecule	GI absorption	BBB permeant	P-gp substrate	CYP inhibitor (Any)*	Bioavailability score	Log K _p (Skin Perm.)
SA	Low	No	Yes	No	0.11	-11.4 cm/s
Analog 5	High	No	No	No	0.56	-8.30 cm/s
Analog 4	High	No	No	No	0.56	-6.87 cm/s
Analog 3	High	Yes	No	No	0.85	-6.66 cm/s

The second phase of the ADME analysis focuses on the physiological transit and metabolic interactions of the compounds as tabulated in Table 7. The results reveal a transformative improvement in the predictive absorption and bioavailability profiles of the AI-derived analogs compared to the natural SA. The most striking distinction lies in the predicted GI absorption. SA is characterized by “Low” absorption and a negligible Bioavailability Score (0.11), largely due to its high molecular weight and excessive polarity. In contrast, all synthetic analogs (3, 4, and 5) exhibit “High” GI absorption. Analog 3 stands out with a Bioavailability Score of 0.85, suggesting nearly optimal oral drug-likeness. Conversely, none of the synthetic analogs are P-gp substrates. The structural tuning of these analogs enables highly specific therapeutic targeting. Notably, Analog 3 is

the only compound predicted to be BBB permeant, suggesting its potential for central nervous system (CNS) applications where crossing the blood-brain barrier is a prerequisite. Furthermore, the skin permeability (log K_p) data indicates a substantial improvement in transdermal potential; the analogs exhibit significantly higher permeability constants than SA, which shifts from a highly resistant -11.4 cm/s to a more permeable -6.66 cm/s in Analog 3. This less negative log K_p value confirms a faster rate of skin penetration, positioning these synthetic scaffolds as superior candidates for transdermal delivery systems compared to the natural parent material. Notably, none of the compounds—including the analogs—act as Cytochrome P450 (CYP) inhibitors. (Cytochrome P450 enzymes are responsible for the biotransformation of the vast majority of clinical drugs in

the liver). This absence of cytochrome P450 inhibition is a strong indicator of a favorable safety profile, as it minimizes the risk of adverse drug-drug interactions (DDIs) when co-administered with other medications.

The BOILED-Egg analysis provides a visual representation of the absorption and brain permeability profiles of SA and its synthetic analogs based on the relationship between lipophilicity (WLOGP) and topological polar surface area (TPSA) as depicted in Fig. 8. SA is positioned far outside both the human intestinal absorption (HIA) region (white) and the blood–brain barrier (BBB) permeation zone (yolk), reflecting its excessively high polarity and poor membrane permeability. This distribution is consistent with its low predicted gastrointestinal absorption and limited systemic bioavailability. In contrast, Analogs 3, 4, and 5 cluster within or near the HIA region, indicating substantially improved oral absorption potential resulting from optimized polarity and balanced lipophilicity. Notably, Analog 3 is located within the BBB-permeable region, suggesting favorable central nervous system accessibility, which aligns with its moderate lipophilicity and reduced polar surface area. Analog 4 remains proximal to the BBB boundary, implying partial brain accessibility with preferential peripheral distribution, whereas Analog 5 resides primarily within the intestinal absorption domain, supporting its suitability for systemic and localized therapeutic delivery without unintended CNS penetration.

Molecular docking

Protein structure modeling and validation

To obtain a complete and reliable receptor for docking, the AURKA catalytic domain (residues 127–389) was modeled with AlphaFold3, restoring the disordered segment

(residues 278–290) absent from the experimental structure 8JF4. The model displayed excellent stereochemical quality across all validation metrics. VERIFY3D indicated that 88.59% of residues attained an averaged 3D–1D score ≥ 0.1 , exceeding the acceptance threshold, while ERRAT returned an overall quality factor of 91.56 (Fig. 9a), indicating a well-resolved, low-error backbone. PROCHECK Ramachandran analysis assigned 94.0% of residues to the most favored regions (Fig. 9c), consistent with a stereochemically sound fold, and the SWISS-MODEL QMEAN4 score of 0.45 (Fig. 9d) confirmed the energetic plausibility of the model. Structural superposition with the native 8JF4 structure gave a backbone RMSD of 0.893 Å (Figs. 9e and b), demonstrating that the modeled domain—including the reconstructed loop—closely reproduces the experimental geometry. The subsequent 50 ns MD simulation confirmed that the model remained structurally stable throughout the trajectory, justifying its use as the docking receptor.

Molecular docking

Molecular docking was performed to evaluate the binding affinity of the co-crystallized ligand (C0N), SA, and the AI-designed synthetic biomimetic analogs (Analog 5, Analog 4, and Analog 3) toward the active site of Aurora kinase A (AURKA). The binding poses and 2D interaction maps are presented in Fig. 10, while the corresponding Glide and Induced Fit Docking (IFD) scores are summarized in Table 8.

The C0N was first analyzed to identify the key interactions within the ATP-binding pocket and to serve as a reference for evaluating the binding modes of the designed ligands. C0N exhibited a Glide score of -5.75 and an IFD score of -9.69 , establishing 2 hydrogen-bond and hydrophobic interactions with residues lining the catalytic pocket (Fig. 10b). Among all investigated compounds, SA exhibited the strongest binding affinity, with a Glide score of -6.142 and an IFD score of -11.23 , indicating the most favorable ligand–protein interaction after accounting for receptor flexibility. The superior binding of SA can be attributed to its polysaccharide backbone containing multiple hydroxyl and carboxylate groups that provide numerous hydrogen-bond donor and acceptor sites. As shown in Fig. 10c, SA established 9 hydrogen-bonding network together with electrostatic and hydrophobic interactions involving residues surrounding the catalytic pocket. Among the AI-designed biomimetic analogs, Analog 5 demonstrated the highest binding affinity, with a Glide score of -5.405 and an IFD score of -7.97 . As illustrated in Fig. 10d, Analog 5 retained 5 hydrogen-bond interactions observed. Its tetrahydropyran scaffold and two carboxylic acid groups effectively mimic the spatial arrangement of the uronic acid units of alginate, allowing

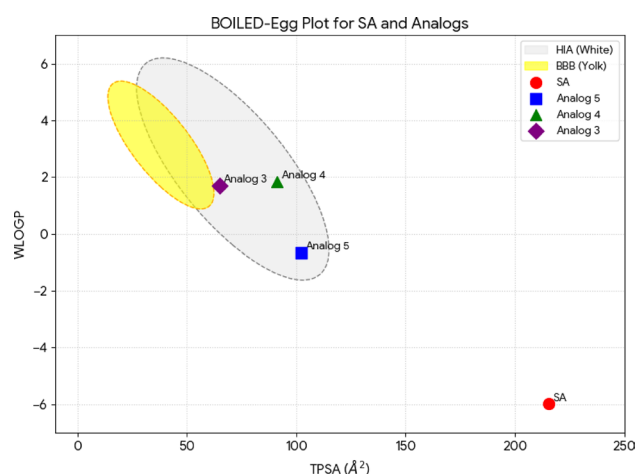


Fig. 8 BOILED-Egg plot (Brain Or Intestinal Estimated permeation method) for SA and its synthetic analogs

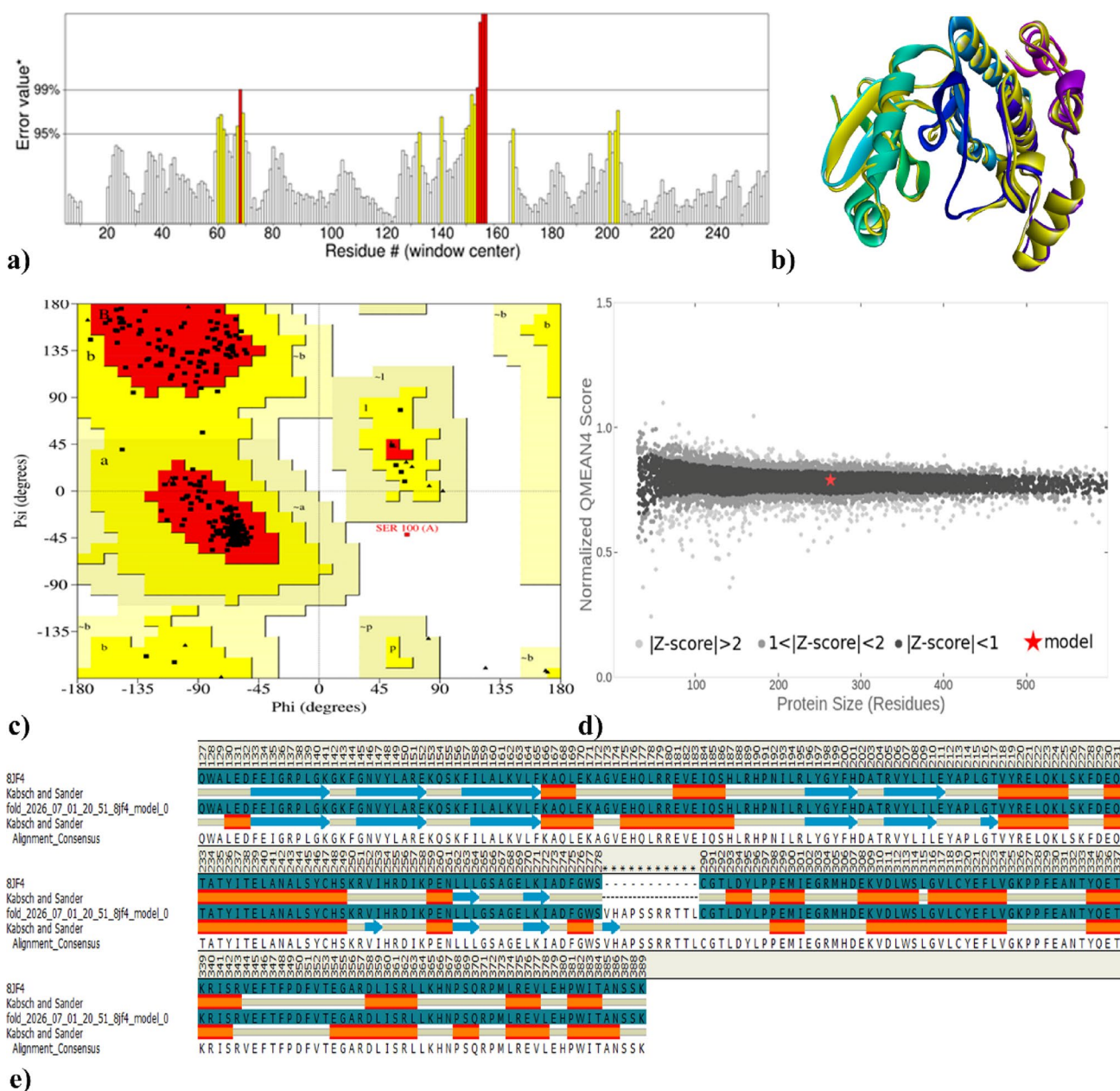


Fig. 9 Modeling and stereochemical validation of the human AURKA catalytic domain (residues 127–389). **a)** ERRAT plot of error value versus residue number (window center), giving an overall quality factor of 91.56. On the error axis, the two horizontal lines mark the 95% and 99% confidence limits above which residue regions may be rejected; bars exceeding these limits indicate misfolded or error-prone segments. Good high-resolution structures generally yield values around 95% or higher, whereas for lower-resolution structures (2.5–3.0 Å) the average overall quality factor is typically around 91%, placing the present model within the expected range for a reliable structure. **b)** Structural superposition of the AlphaFold3 model onto the experimental structure 8JF4, showing close backbone agreement (RMSD=0.893 Å); the native structure with missing residues is shown in yellow and the model in ribbon colour. **c)** Ramachandran plot generated by PROCHECK. Of the 233 non-glycine, non-proline residues,

219 (94.0%) lie in the most favoured regions [A, B, L], 13 (5.6%) in additional allowed regions [a, b, I, p], 0 (0.0%) in generously allowed regions [~a, ~b, ~I, ~p], and 1 (0.4%) in disallowed regions. The model also contained 2 end-residues, 15 glycine residues (triangles), and 13 proline residues, for a total of 263 residues. Based on an analysis of 118 structures at ≥ 2.0 Å resolution and $R\text{-factor} \leq 20\%$, a good-quality model is expected to have over 90% of residues in the most favoured regions, a criterion satisfied here. **d)** Normalized QMEAN4 score of the model (red star) plotted against protein size relative to a non-redundant set of PDB structures, confirming its energetic reliability (QMEAN4=0.45). **e)** Residue-level secondary-structure alignment (DSSP/Kabsch-Sander) between 8JF4 and the AlphaFold3 model, illustrating conservation of the secondary-structure elements and reconstruction of the disordered loop (residues 278–290) absent from the experimental structure

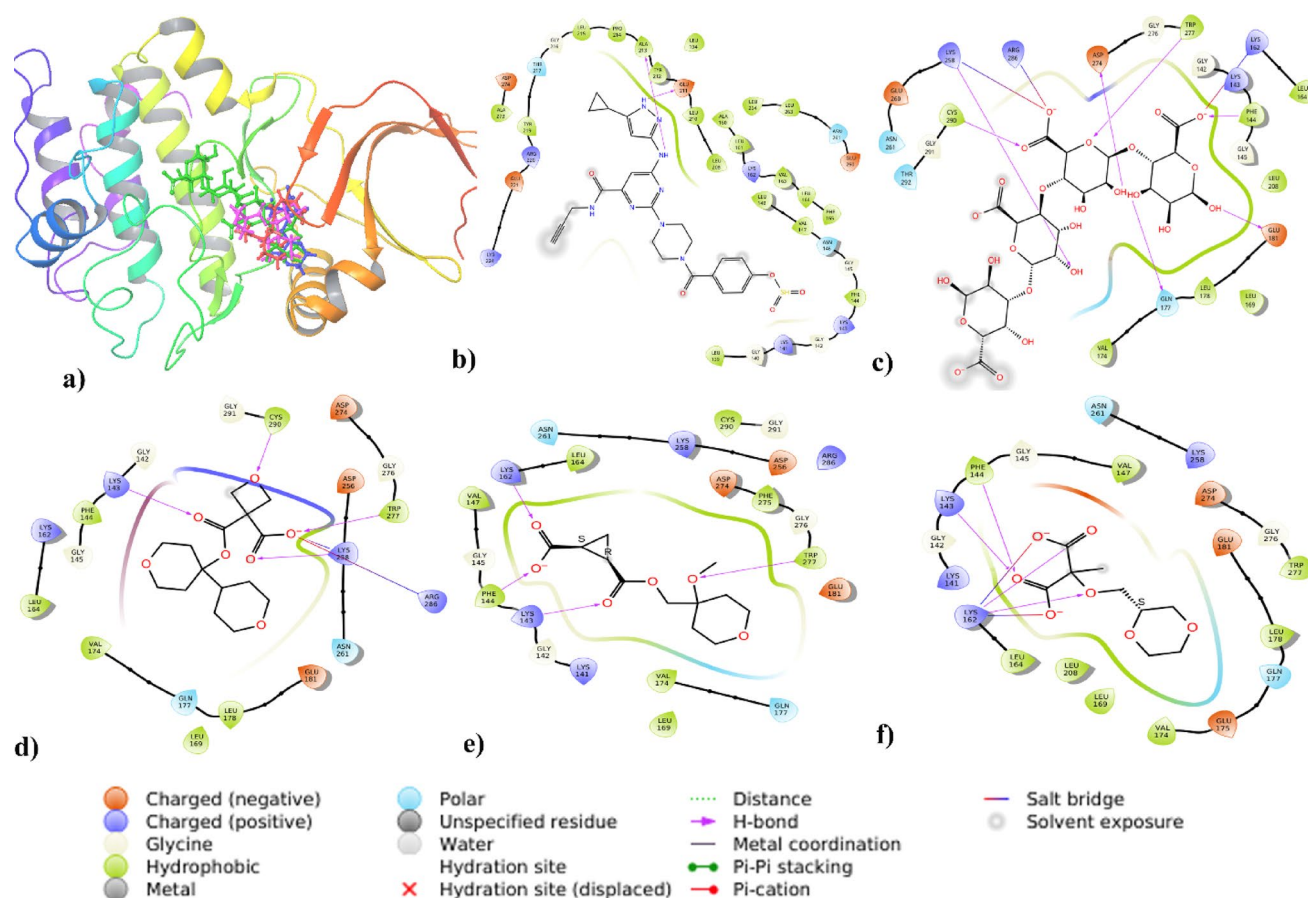


Fig. 10 Molecular docking analysis of the co-crystallized ligand, SA, and AI-designed synthetic biomimetic analogs within the active site of AURKA. **a)** 3D representation of the binding poses of the co-crystallized ligand and the docked ligands within the AURKA active site.

Table 8 Molecular docking results of SA and AI-designed synthetic analogs against the target protein, showing the Glide docking scores and Induced Fit Docking (IFD) scores

Ligands	Glide Score	Induced Fit Docking Score
CON	− 5.75	− 9.69
SA	− 6.142	− 11.23
Analog_5	− 5.405	− 7.97
Analog_4	− 4.496	− 7.77
Analog_3	− 4.147	− 7.10

favorable polar interactions with active-site residues. Analog 4 exhibited intermediate binding affinity with a Glide score of −4.496 and an IFD score of −7.77 (Fig. 10e). The ligand formed 4 hydrogen bonds primarily through its ester and carboxyl oxygen atoms, while additional van der Waals and hydrophobic interactions contributed to stabilization of the complex. Similarly, Analog 3 produced a Glide score of −4.147 and an IFD score of −7.10 (Fig. 10f), representing the weakest interaction among the designed analogs. Nevertheless, Analog 3 adopted a well-defined binding orientation

2D-dimensional interaction diagram of **b)** the co-crystallized ligand SA, **d)** Analog 5, **e)** Analog 4, and **f)** Analog 3 showing their intermolecular interactions with the active-site residues

and maintained 6 hydrogen-bond interactions with residues in the catalytic pocket.

The molecular docking results are consistent with the quantum chemical and physicochemical analyses presented earlier. Native SA exhibited the strongest binding affinity toward the target protein, with a Glide score of −6.142 and an Induced Fit Docking (IFD) score of −11.23, which can be attributed to its highly polar structure, abundant hydroxyl and carboxylate groups, and extensive hydrogen-bonding capability. These characteristics are further supported by its high topological polar surface area (TPSA, 215.50 Å²), 13 hydrogen-bond acceptors, 6 hydrogen-bond donors, consensus LogP of −4.18, and the highest electronic softness (0.123 eV^{−1}) together with an electrophilicity index of 2.031 eV, facilitating favorable electrostatic interactions within the binding pocket.

The AI-designed synthetic analogs were developed to preserve the essential structural features responsible for alginate recognition while substantially improving drug-like properties. Analog 5 showed the highest affinity among the designed compounds, with a Glide score of −5.405 and

an IFD score of -7.97 , indicating that the tetrahydropyran scaffold containing two carboxylic acid groups effectively mimics the uronic acid units of alginate. Compared with SA, Analog 5 exhibited a significantly lower molecular weight ($234.20 \text{ g mol}^{-1}$), reduced TPSA ($102.29 \text{ \AA}^2$), zero Lipinski violations, high gastrointestinal absorption, and an improved bioavailability score (0.56), while maintaining strong protein-binding capability.

Analog 4 demonstrated intermediate binding (Glide score = -4.496 ; IFD score = -7.77) together with favorable physicochemical characteristics, including a molecular weight of $310.30 \text{ g mol}^{-1}$, TPSA of $91.29 \text{ \AA}^2$, consensus LogP of 1.77 , high GI absorption, and zero Lipinski violations, suggesting a balanced combination of binding affinity and pharmacokinetic performance. Analog 3 exhibited the lowest docking affinity (Glide score = -4.147 ; IFD score = -7.10), yet possessed the most favorable predicted pharmacokinetic profile, with the lowest TPSA ($64.99 \text{ \AA}^2$), molecular weight of $252.26 \text{ g mol}^{-1}$, the highest bioavailability score (0.85), high gastrointestinal absorption, and predicted blood–brain barrier permeability, making it an attractive scaffold for applications requiring enhanced systemic or central nervous system delivery.

Molecular dynamic simulation

To further evaluate the structural stability of the docked complexes under physiological conditions, 300 ns molecular dynamics (MD) simulations were performed for the SA-, Analog 5-, Analog 4-, and Analog 3-bound protein complexes. The temporal evolution of the protein backbone root mean square deviation (RMSD) and the residue-wise root mean square fluctuation (RMSF) are presented in Fig. 11(a and b).

The RMSD profiles (Fig. 11a) demonstrate that all ligand–protein complexes rapidly reached equilibrium within the initial 20–30 ns, indicating successful relaxation of the docked conformations. Throughout the 300 ns simulation, the SA complex exhibited a relatively constant RMSD ranging between 0.28 and 0.34 nm, among the synthetic analogs, Analog 5 displayed the highest overall stability, maintaining RMSD values predominantly between 0.24 and 0.30 nm throughout the simulation. In contrast, the Analog 4 complex exhibited transient increases in RMSD between approximately 150 ns and 200 ns, reaching values close to 0.55–0.60 nm.

The residue-wise RMSF analysis (Fig. 11b) provides insight into the local flexibility of the protein in the presence of different ligands. For all three analyzed complexes (SA, Analog 5, and Analog 3), the majority of amino acid residues exhibited fluctuations below 0.20 nm, indicating a generally rigid protein framework during simulation.

Higher fluctuations were observed mainly at the N-terminal and C-terminal regions, together with several solvent-exposed loop segments around residues approximately 150, 280–290, 330–340, and the terminal residues, where RMSF values increased to approximately 0.30–0.55 nm. These flexible regions are characteristic of loop and terminal segments and are commonly observed in long-timescale MD simulations.

Importantly, the binding-site residues displayed comparatively low fluctuations throughout the simulation, suggesting that ligand binding effectively stabilized the catalytic pocket. The RMSF profiles of SA, Analog 5, and Analog 3 largely overlapped, indicating that the synthetic biomimetic analogs preserve the dynamic behavior of the native polymer–protein complex. Although Analog 5 exhibited slightly higher fluctuations at several loop regions, the differences were minimal and did not affect the stability of the active-site architecture.

Collectively, the RMSF results indicate that ligand binding does not induce significant structural perturbations in the protein and that the observed flexibility is primarily restricted to peripheral loop regions rather than residues directly involved in ligand recognition.

MM-PBSA binding energies

To quantitatively evaluate the binding affinity and energetic stability of the ligand–protein complexes beyond molecular docking, the Molecular Mechanics Poisson–Boltzmann Surface Area (MM-PBSA) method was employed. A total of 1000 snapshots extracted from the equilibrated region of the 300 ns molecular dynamics trajectory (280–290 ns) at 10 ps intervals were used for the free energy calculations. The total binding free energy (BE) was decomposed into van der Waals (vdW-E), electrostatic (ESE), polar solvation (PSE), and solvent-accessible surface area (SASA-E) energy components, and the calculated values are summarized in Table 9.

The MM-PBSA results reveal that all three ligands exhibit negative binding free energies, indicating thermodynamically favorable complex formation with the target protein. Among the investigated ligands, SA exhibited the strongest binding affinity, with a total binding free energy of $-666.260 \pm 1.291 \text{ kJ mol}^{-1}$. This exceptionally favorable binding is primarily driven by a remarkably large electrostatic contribution ($-1529.220 \text{ kJ mol}^{-1}$), reflecting the strong ionic interactions between the multiple negatively charged carboxylate groups of SA and complementary positively charged residues within the protein binding pocket. In addition, favorable van der Waals interactions ($-60.253 \text{ kJ mol}^{-1}$) and nonpolar solvation energy (SASA-E = $-16.472 \text{ kJ mol}^{-1}$) further stabilize the complex.

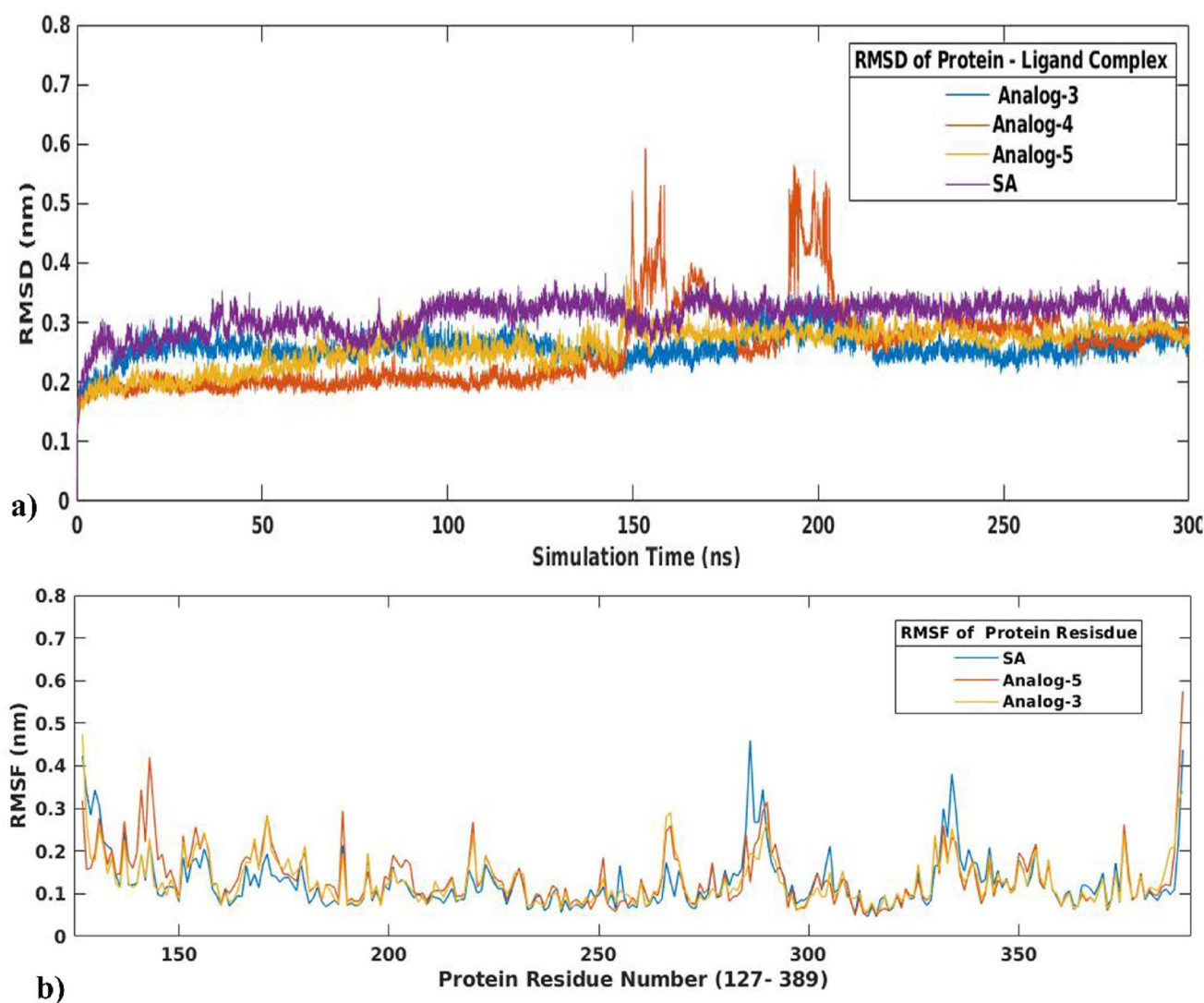


Fig. 11 MD simulation analysis of the protein complexes bound with SA and the AI-designed synthetic biomimetic analogs. **a)** RMSD of the protein backbone during the 300 ns MD simulation for the SA, Analog 5, Analog 4, and Analog 3 complexes, demonstrating the structural

stability of the ligand–protein complexes over time. **b)** RMSF of individual amino acid residues (residues 127–389) for the SA, Analog 5, and Analog 3 complexes, illustrating residue-wise flexibility throughout the simulation

Table 9 MM-PBSA binding free energy decomposition of SA and the AI-designed synthetic biomimetic analogs in complex with the target protein, calculated from 1000 snapshots extracted from the equilibrated 300 ns molecular dynamics trajectory (280–290 ns) at 10 ps intervals

Sl.No	Ligand	vdW-E (kJ mol ⁻¹)	ESE (kJ mol ⁻¹)	PSE (kJ mol ⁻¹)	SASA-E (kJ mol ⁻¹)	BE (kJ mol ⁻¹)	SD
1	SA	− 60.253	− 1529.220	939.710	− 16.472	− 666.260	1.291
2	Analog_5	− 84.034	− 18.558	64.287	− 9.421	− 47.735	0.315
3	Analog_3	− 138.645	− 301.725	223.654	− 15.283	− 232.137	0.772

Although the polar solvation energy (+939.710 kJ mol⁻¹) opposes binding due to the energetic cost of desolvating the highly polar ligand, it is insufficient to overcome the strong electrostatic stabilization, resulting in the most favorable overall binding energy among the studied compounds.

Among the synthetic biomimetic candidates, Analog 3 demonstrated substantially stronger binding than Analog 5, exhibiting a total binding free energy of -232.137 ± 0.772 kJ

mol⁻¹. Unlike SA, the binding of Analog 3 is governed by a balanced combination of strong van der Waals interactions (-138.645 kJ mol⁻¹) and electrostatic interactions (-301.725 kJ mol⁻¹). The relatively lower polar solvation penalty ($+223.654$ kJ mol⁻¹) compared with SA reflects the reduced polarity of the synthetic scaffold, allowing the favorable intermolecular interactions to dominate the overall binding process. The favorable SASA contribution

($-15.283 \text{ kJ mol}^{-1}$) further supports stable complex formation throughout the molecular dynamics simulation.

In contrast, Analog 5 exhibited the least favorable binding free energy ($-47.735 \pm 0.315 \text{ kJ mol}^{-1}$). Although it possessed appreciable van der Waals stabilization ($-84.034 \text{ kJ mol}^{-1}$), its comparatively weak electrostatic interaction ($-18.558 \text{ kJ mol}^{-1}$), together with the opposing polar solvation energy ($+64.287 \text{ kJ mol}^{-1}$), substantially reduced the overall binding affinity. Nevertheless, the negative total binding energy indicates that the complex remains thermodynamically stable, which is consistent with the stable RMSD profile observed during the 300 ns molecular dynamics simulation. The ranking of binding free energies obtained from MM-PBSA follows the order:

SA ($-666.260 \text{ kJ mol}^{-1}$) > Analog 3 ($-232.137 \text{ kJ mol}^{-1}$) > Analog 5 ($-47.735 \text{ kJ mol}^{-1}$).

This trend generally agrees with the molecular docking results, confirming that SA possesses the strongest intrinsic binding affinity toward the target protein. Importantly, despite their reduced binding energies, the synthetic analogs—particularly Analog 3—maintain favorable thermodynamic stability while offering significantly improved pharmacokinetic and drug-like properties.

Conclusions

DFT calculations revealed intrinsic stability differences between alginate monomers, with β -D-mannuronic acid exhibiting slightly lower energy ($-1,998,257.29 \text{ kJ mol}^{-1}$) than α -L-guluronic acid ($-1,998,251.23 \text{ kJ mol}^{-1}$). Among the dimers, the MG block ($-3,795,894.49 \text{ kJ mol}^{-1}$) was more stable than the MM block ($-3,795,878.96 \text{ kJ mol}^{-1}$), while progressive polymerization further stabilized the structure, with the MGMMGG hexamer reaching $-10,986,523.52 \text{ kJ mol}^{-1}$. Sodium coordination significantly enhanced stability, yielding a total energy of $-9,089,755.72 \text{ kJ mol}^{-1}$. Frontier molecular orbital analysis showed a decrease in the HOMO–LUMO gap from 10.75 to 10.96 eV (monomers) to 10.27 eV (oligomers) and 8.15 eV (SA), accompanied by reduced chemical hardness (4.07 eV), increased softness (0.123 eV^{-1}), and an electrophilicity index of 2.031 eV, indicating enhanced electronic delocalization and interaction capability. The optimized model was validated by the close agreement between theoretical (1646 and 1461 cm^{-1}) and experimental (1641 and 1413 cm^{-1}) FTIR spectra.

AI-guided molecular design produced synthetic alginate analogs with markedly improved drug-like properties. Compared with SA (355.23 g mol^{-1} , TPSA 215.50 $\text{\AA}^2$, bioavailability 0.11), the analogs showed reduced molecular weights (234.20–310.30 g mol^{-1}), lower TPSA (64.99–102.29 $\text{\AA}^2$), zero Lipinski violations, and higher predicted

bioavailability (0.56–0.85). Analog 3 additionally exhibited predicted blood–brain barrier permeability and all analogs showed improved skin permeability.

Molecular docking against Aurora kinase A (AURKA) demonstrated the strongest binding for SA (Glide score = -6.142 ; IFD score = -11.23), followed by Analog 5 (-5.405 ; -7.97), Analog 4 (-4.496 ; -7.77), and Analog 3 (-4.147 ; -7.10). 300 ns molecular dynamics simulations confirmed stable ligand–protein complexes with only minor loop fluctuations. MM-PBSA calculations further supported the docking results, with binding free energies of $-666.260 \pm 1.291 \text{ kJ mol}^{-1}$ for SA, $-232.137 \pm 0.772 \text{ kJ mol}^{-1}$ for Analog 3, and $-47.735 \pm 0.315 \text{ kJ mol}^{-1}$ for Analog 5, indicating that electrostatic interactions dominate SA binding, whereas the synthetic analogs rely on a balanced contribution of van der Waals and electrostatic interactions.

Overall, the integration of DFT, FTIR validation, AI-guided molecular design, ADME prediction, molecular docking, molecular dynamics simulation, and MM-PBSA analysis demonstrates that the designed biomimetic analogs successfully combine favorable binding characteristics with significantly improved pharmacokinetic properties. Analog 5 emerged as the closest functional mimic of SA with excellent dynamic stability, whereas Analog 3 provided the most favorable balance between binding free energy and drug-likeness, making both promising candidates for next-generation alginate-inspired biomaterials for drug delivery, wound healing, and tissue engineering. Future studies will investigate explicit hydration effects using computational modeling coupled with broadband dielectric spectroscopy to better understand alginate behavior under physiological conditions.

Author contributions Safna Hussan K.P: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – original draft. Sooraj Kumar S: Data curation, Validation, Visualization. Jibin K Varghese- MD simulation part, MM-PBSA binding free energy, Rio Kita: Resources, Supervision, Writing – review & editing. Naoki Shinyashiki: Project administration, Supervision, Writing – review & editing.

Funding Open Access funding provided by Tokai University. The authors declare that no external funding was received for this research.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the

Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Hurtado A, Aljabali AAA, Mishra V, Tambuwala MM, Serrano-roca Á (2022) Alginate: enhancement strategies for advanced applications. *Int J Mol Sci Rev* 23:1–45
- Abka-khajouei R, Tounsi L, Shahabi N, Patel AK, Abdelkafi S (2022) Structures, properties and applications of alginates. *Mar Drugs* 20:1–18
- Seto D, Anugrah B, Darmalim LV, Rifky M, Polanen I, Putro PA, Sasongko NA, Siahaan P, Ramadhan ZR (2022) Quantum chemical calculation for intermolecular interactions of alginate dimer-water molecules. *Gels* 8(11), 703
- Mineva T (2012) Structure of alginate gels: interaction of diuronate units with divalent cations from density functional calculations. *Biomacromolecules* 13:1899–1907. <https://doi.org/10.1021/bm300420z>
- Zhang M, Hong H, Lin H, Shen L, Yu H, Ma G & Liao BQ (2018) Mechanistic insights into alginate fouling caused by calcium ions based on terahertz time-domain spectra analyses and DFT calculations. *Water Res* 129: 337–346.
- Arooj M, Thangapandian S, John S, Hwang S, Park JK, & Lee KW (2012) Computational studies of novel chymase inhibitors against cardiovascular and allergic diseases: mechanism and inhibition. *Chem Biol Drug Des* 80:862–875
- Nemat AAM, AlBab D, Mehavesh K, Hameed A, Maresova IM, Ahmady M, Arooj C, Han B, Workie M, Chehimi (2020) Inhibition of amyloid fibrillation, enzymatic degradation and cytotoxicity of insulin at carboxyl tailored gold-aryl nanoparticles surface. *Colloids Surf. Physicochem Eng Asp* 586:124279
- Tangella LP, Arooj M, Deplazes E, Gray ES, Mancera RL (2021) Identification and characterisation of putative drug binding sites in human ATP-binding cassette B5 (ABCB5) transporter. *Comput. Struct Biotechnol J* 19:691–704. <https://doi.org/10.1016/j.csbj.2020.12.042>
- Fahmy A, Khafagy RM, Elhaes H, Ibrahim MA, Spectroscopy M, Unit M (2021) Egypt J Chem 64:1149–1166. <https://doi.org/10.21608/EJCHEM.2021.55865.3184>
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Scalmani CJRG, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven ET, Montgomery JA Jr., Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers A, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell MK, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Knox RESJE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Yazyev KMO, Austin AJ, Cammi R, Pomelli C, Ochterski J, Martin RL, Zakrzewski OVG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas DF, Foresman JB, Ortiz JV, Cioslowski J (1998) J. GAUSSIAN 09 (Revision A.2) Gaussian, Inc., Wallingford, CT, D. J. Gaussian 09, B.01
- Head-Gordon M (2008) Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys Chem Chem Phys* 10(44):6615–6620
- Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys Chem Chem Phys* 7:3297–3305
- Reidenbach D, Livne M, Ilango RK, Gill M, Israeli J (2023) Improving Small Molecule Generation using Mutual Inform Mach. <http://arxiv.org/abs/2208.09016>
- Livne M, Swersky K, Fleet DJ (2020) MIM: Mutual Inform Mach. <http://arxiv.org/abs/1910.03175>
- Manual U Glide 5.5, (n.d.)
- Hussan S Influence of intermolecular hydrogen bonding on the dynamics of water and polymer in PEI water mixtures influence of intermolecular hydrogen bonding on the dynamics of water and polymer in pei water mixtures T g [K], (n.d.).
- Ertl P, Schuffenhauer A (2009) Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J Cheminform* 1:1–11. <https://doi.org/10.1186/1758-2946-1-8>
- SWISSADME (n.d.)
- Prof.Dr SQY, Dr. Guanghui Tang W, Wang C, Zhu H, Huang P, Chen X, Wang M, Xu J, Sun D, Chong-Jing, Zhang (2024) Dr. Qicai Xiao, Dr. Liqian Gao, Dr. Zhi-Min Zhang, Global reactivity profiling of the catalytic lysine in human kinome for covalent inhibitor development. *Angew Chem Int Ed* 63:e202316394
- Eisenberg D, Lüthy R, Bowie JU, VERIFY3D (1997) Assessment of protein models with three-dimensional profiles. *Methods Enzymol* 277:396–404. [https://doi.org/10.1016/S0076-6879\(97\)77022-8](https://doi.org/10.1016/S0076-6879(97)77022-8)
- Colovos YT (1993) Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci* 2:1511–1519
- Andrew Waterhouse FT, Bienert MBS, Studer G, Tauriello G, Gumienny R (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46:w296–w303
- Ramachandran SV, Ramakrishnan GN (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95–99
- BIOVIA (2020) Materials Studio, Release 2020 (or your version). Dassault Systèmes, San Diego, CA, USA.
- Schrödinger N (2021) LLC.New York, LigPrep Schrödinger Release 2021-2
- Glide N, Schrödinger LLC (2021) New York, Schrödinger Release 2021-2
- Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2014) Glide: A new approach for rapid, accurate docking and scoring. 2. enrichment factors in database screening. *J Med Chem* 47:1750–1759
- van der Berendsen D (1995) GROMACS: A message-passing parallel molecular dynamics implementation, 91(1–3):43–56. [https://doi.org/https://doi.org/10.1016/0010-4655\(95\)00042-E](https://doi.org/https://doi.org/10.1016/0010-4655(95)00042-E)
- Tirado-Rives WLJ;DSMJ (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236
- Zoete MO, Cuendet V, Grosdidier MA (2011) SwissParam: a fast force field generation tool for small organic molecules. *J Comput Chem* 32:2359–2368
- Bugnon ZV, Goullieux M, Röhrig M, Perez UF, Daina MAS, Michielin A, Bugnon O, Goullieux M, Röhrig M, Perez UF, Daina MAS, Michielin A, Zoete O (2023) V. SwissParam 2023: a modern web-based tool for efficient small molecule parameterization. *J Chem Inf Model* 63(21): 6469–6475
- Leela WLJ, Dodda S (2017) Israel Cabeza de Vaca, Julian Tirado-Rives, LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res* 45:W331–W336. <https://doi.org/10.1093/nar/gkx312>

33. Toukan K, Rahman A (1985) Molecular-dynamics study of atomic motions in water. *Phys Rev B* 31(5):2643
34. Berendsen JR, Postma HJC, van Gunsteren JPM, DiNola WF (1984) Haak, Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690
35. Michele Di BL, Pierro R, Elber (2015) A Stochastic Algorithm for the Isobaric-Isothermal Ensemble with Ewald Summations for All Long Range Forces. *J Chem Theory Comput* 11:5624–5637. <https://doi.org/10.1021/acs.jctc.5b00648>
36. Scheurer RCBM, Rodenkirch P, Siggel M, Schulten ETK, Rudack T (2018) PyContact: rapid, customizable, and visual analysis of noncovalent interactions in MD simulations. *Biophys J* 114(3):577–583. <https://doi.org/10.1016/j.bpj.2017.12.003>
37. Kollman TECPA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* 33:889–897
38. Zoete OMV, Irving MB (2010) MM–GBSA binding free energy decomposition and T cell receptor engineering. *J Mol Recognit* 23:142–152
39. Sattar F, Wang Z, Zhou X, Ullah Z (2024) Battling hazardous gas molecules with kekulene surfaces: A computational study. *J Mol Liq* 407:125099. <https://doi.org/10.1016/j.molliq.2024.125099>
40. Ullah Z, Kim HJ, Sheena Mary Y, Kwon HW (2022) Insights into caffeine adsorption on the surface of corannulene: A sensor study. *J Mol Liq* 368:120592. <https://doi.org/10.1016/j.molliq.2022.120592>
41. Ullah Z, Jee Kim H, Sheena Mary Y, Belboukhari N, Sekkoum K, Kraimi A, Zhan X, Wook H, Kwon (2023) Unlocking the potential of ovalene: A dual-purpose sensor and drug enhancer. *J Mol Liq* 377:121540. <https://doi.org/10.1016/j.molliq.2023.121540>
42. Kar R, Chandrakumar KRS, Pal S (2007) The Influence of Electric Field on the Global and Local Reactivity Descriptors: Reactivity and Stability of Weakly Bonded Complexes. *J Phys Chem A* 111:375–383
43. Babu NS, Jayaprakash D (2015) Global and reactivity descriptors studies of cyanuric acid tautomers in different solvents by using of density functional theory (DFT), 4 615–620
44. Sachdeva R, Soni A, Singh VP, Saini GSS (2018) Reactivity of etoricoxib based on computational study of molecular orbitals, molecular electrostatic potential surface and Mulliken charge analysis. *AIP Conf Proc* 1953:1–5. <https://doi.org/10.1063/1.5033181>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.