International Journal of Oral Science www.nature.com/ijos

ARTICLE OPEN

Machine learning-driven discovery of therapeutic nucleoside hydrogels for periodontitis

Weiqi Li[1,2], Yinghui Wen[1], Zhenyuan Huang[1], Fangyuan Shuai[1], Yijia Yin[1], Qianming Chen[1], Fudong Zhu[3], Hao Xu[1]✉ and Hang Zhao[1]✉

Supramolecular hydrogels hold significant potential in drug delivery and tissue engineering, with standing out for their unique properties. Despite their promise, predicting nucleoside bioactivity remains challenging. This study aims to predict the biological activity of nucleosides to guide the rational synthesis of hydrogels. Specifically, nine predictive models and databases for various biological activities were built with feature-selected machine learning methods including decision trees, logistic regression, random forest, and extreme gradient boosting. Then, the Molecular Bioactivity Specificity Index (MBSI) was introduced to gauge the primary bioactivity of nucleoside derivatives, and the Composite Molecular Attribute Score (CMAS) was devised to measure the overall performance of nucleoside derivatives. Subsequently, screening strategies for bioactive nucleoside hydrogels were established, and two candidate hydrogels (GMP and dGMP) with high hydrogel-forming ability, biocompatibility, and antibacterial activity were identified. Finally, two hydrogels were validated for antibacterial treatment of periodontitis. This study highlights the feasibility of ML-based strategies and MBSI/CMAS in rationally designing bioactive nucleoside hydrogels for biomedical applications. The discovery of GMP and dGMP hydrogels and their successful validation in periodontitis models highlight the potential of this strategy for developing targeted therapies for oral diseases.

International Journal of Oral Science (2026)18:41 ; https://doi.org/10.1038/s41368-026-00438-3

INTRODUCTION

Supramolecular hydrogels formed through non-covalent interactions possess tunable mechanical properties and responsiveness to external stimuli, making them highly suitable for applications in drug delivery and tissue engineering.[1,2] Among these, nucleoside-based hydrogels have attracted considerable attention due to their remarkable biocompatibility, bioactivity, and self-healing capabilities.[3–7] However, a key challenge in the field of nucleoside hydrogels is that the existing bioactive hydrogels are often discovered by chance or through empirical modification of nucleoside functional groups. A systematic framework for guiding the design and development of hydrogels with specific functionalities is absent. Currently, two primary issues are accurately predicting the gelation potential of nucleoside derivatives, and elucidating the biological functions of these prospective hydrogels.[8–10]

Machine learning (ML) and other artificial intelligence (AI) methods have been widely adopted for constructing predictive models of compound properties and hold significant potential for addressing critical challenges in the field of nucleoside hydrogels.[10–12] ML methodologies can process large datasets and uncover structure-property relationships through high-dimensional data analysis, thereby exploring complex relationships between molecular structures and properties.[10,13] This data-driven approach allows for the rapid screening of candidate compounds with specific targets, reducing experimental costs and shortening development timelines, providing a scientific basis for precise compound synthesis.[10,11] For instance, Guo et. al. used ML and big data mining to successfully predict gas diffusion rates in metal-organic frameworks. Peteani et. al. utilized ML models to design targeted protein degraders, potentially accelerating drug discovery.[13] Stokes et. al. applied ML algorithms to discover new antibiotics effective against *Escherichia coli* and *Acinetobacter baumannii*.[14] In previous studies, our research team has focused on ML to construct predictive models, achieving progress in the first challenge. Based on molecular descriptors, our team developed an ML model to predict the gelation capacity of nucleoside derivatives with an accuracy of 71%, achieving an accuracy of 83.3% in external validations.[10] Building upon this, we aim to apply ML methodologies to further predict the bioactivity of nucleoside hydrogels. However, previous studies have predominantly focused on individual or isolated properties, presenting challenges in predicting multiple biological functions and screening nucleoside derivatives with potential bioactivity.[10,11,14] For instance, many compounds may exhibit excellent antibacterial activity while performing poorly in terms of biocompatibility. Compared to models that rely on single or independent factors, employing systematic screening criteria to evaluate and balance

[1]State Key Laboratory of Oral Diseases & National Center for Stomatology & National Clinical Research Center for Oral Diseases & Research Unit of Oral Carcinogenesis and Management & Chinese Academy of Medical Sciences, West China Hospital of Stomatology, Sichuan University, Chengdu, China; [2]Department of Periodontics, West China Hospital of Stomatology, Sichuan University, Chengdu, China and [3]Stomatology Hospital, School of Stomatology, Zhejiang University School of Medicine, Zhejiang Provincial Clinical Research Center for Oral Diseases, Key Laboratory of Oral Biomedical Research of Zhejiang Province, Cancer Center of Zhejiang University, Engineering Research Center of Oral Biomaterials and Devices of Zhejiang Province, Hangzhou, China
Correspondence: Hao Xu (hao.xu@scu.edu.cn) or Hang Zhao (zhaohangahy@scu.edu.cn)
These authors contributed equally: Weiqi Li, Yinghui Wen, Zhenyuan Huang

Received: 13 June 2025 Revised: 31 January 2026 Accepted: 9 February 2026
Published online: 11 May 2026

various indicators is crucial for predicting the multifactorial bioactivity of nucleoside hydrogels, which in turn helps identify potential target nucleoside derivatives.

To address this need and extend our previous gelation-focused framework, this study incorporates bioactivity prediction into the design pipeline, enabling a function-oriented strategy for nucleoside hydrogel development. It utilizes ML algorithms to construct predictive models for various bioactivities of nucleosides, establish a screening strategy based on ML models and MBSI/CMAS methods for nucleoside derivatives, and design supramolecular nucleoside hydrogels with targeted biofunctions (Supplementary Fig. 1). By combining screening strategy with experimental validation, we strive to develop novel nucleoside hydrogels that meet specific biomedical requirements, with a focus on antibacterial activity for periodontitis treatment.

RESULTS

Development of machine learning models for bioactivity prediction of nucleoside derivatives

Nine bioactivity datasets were collected from PubChem[15] and DeepChem[16] for model construction, including biological toxicology (DeepChem, Tox21),[17] antibacterial activity (PubChem, inhibitors of *Porphyromonas gingivalis*, *P. gingivalis*), anti-viral activities (PubChem, inhibitors of HSV-1, VZV, EBV, and HIV), anti-inflammatory activities (PubChem, inhibitors of IL-1β and TNF-α), and anti-tumor activities (PubChem, inhibitors of oral cancer cell line CAL 27). Detailed information regarding the datasets can be found in Supplementary Method 1 and Supplementary Data 1 - 2. A total of 5,666 molecular descriptors were used as features for model construction (Supplementary Data 3).[10] After excluding missing values, a three-step feature selection process was applied to the molecular descriptors to prevent overfitting and improve model accuracy (Fig. 1a).[18,19] The feature selection steps included rank-sum test, Spearman correlation analysis, and recursive feature elimination (RFE) based on ML algorithms (Supplementary Data 4). To construct comprehensive predictive models, four datasets obtained from the feature selection process were used to build models based on four different ML algorithms (descriptors - Original, descriptors - Wilcoxon, descriptors - Spearman, descriptors - RFE, Fig. 1a-d, Supplementary Fig. 2), including decision tree (DT), logistic regression (LR), random forest (RF), and extreme gradient boosting (XGBoost). And the performance of the three-step feature selection method in this study was also compared to the widely used Least Absolute Shrinkage and Selection Operator (LASSO), the results show that our method is superior (Supplementary Data 5, 6). Details regarding the models are provided in Supplementary Method and Supplementary Table 1. The optimal predictive model was selected based primarily on the test accuracy and Area Under Curve (AUC, Supplementary Method).[10]

Optimal model selection and construction of a bioactive prediction database

The optimal model was determined based on test accuracy and AUC, while also considering recall, f1 score, and precision. Three LR models and six XGBoost models based on RFE were selected as the optimal models for above nine bioactivities predictions (Fig. 1b-d, Fig. 2a-d, Supplementary Figs. 2–4, Supplementary Table 2–10). The optimal model for each bioactivity is as follows: Biological toxicity, XGBoost (Mean ± SE, test accuracy, 0.777 ± 0.001; AUC, 0.785 ± 0.002); antibacterial (*P. gingivalis*), LR (test accuracy, 0.897 ± 0.007; AUC, 0.923 ± 0.008); anti-viral (HSV-1), LR (test accuracy, 0.675 ± 0.001; AUC, 0.551 ± 0.002); anti-viral (VZV), XGBoost (test accuracy, 0.852 ± 0.002; AUC, 0.885 ± 0.003); antiviral (EBV), XGBoost (test accuracy, 0.738 ± 0.004; AUC, 0.806 ± 0.004); anti-viral (HIV), XGBoost (test accuracy, 0.873 ± 0.002; AUC, 0.945 ± 0.001); anti-inflammatory (IL-1β), XGBoost (test accuracy, 0.709 ± 0.004; AUC, 0.779 ± 0.004); anti-inflammatory (TNF-α), XGBoost (test accuracy, 0.833 ± 0.001; AUC, 0.926 ± 0.001); anti-tumor (CAL 27), LR (test accuracy, 0.755 ± 0.008; AUC, 0.790 ± 0.008), detail could be found in Supplementary Table 11. By constructing these bioactivity prediction models, we found that ML models were effective in predicting multiple bioactivities of compounds. Using these nine predictive models, we screened a dataset of 7257 nucleoside derivatives identified in our previous hydrogel-forming ability prediction study[10] (Fig. 2 a). Thereby facilitating the identification of target bioactive nucleoside hydrogels, and further refined our database (NBPND, Novel Bioactive Prediction of Nucleoside Derivatives Database, Supplementary Data 7).

Screening strategy for antibacterial nucleoside hydrogels for periodontitis treatment

Periodontitis is an inflammatory disease that damages periodontal tissues (gingiva, alveolar bone, and periodontal ligament).[20,21] It is considered the sixth most common disease globally, imposing a significant health and economic burden on individuals, families, and society.[22,23] Given the open, moist, and dynamic environment of the oral cavity, our team has previously developed a series of biomimetic nucleoside hydrogel delivery systems for the prevention and treatment of oral diseases.[4–6] However, rationally designing and synthesizing nucleoside hydrogels with customized bioactivities for periodontal treatment remains a major challenge. An ideal therapeutic strategy requires antibacterial properties to inhibit periodontal pathogens, primarily *P. gingivalis*, while ensuring good biocompatibility to minimize inflammatory responses.[24,25] Moreover, the gelation capacity of hydrogels in periodontal tissues affects their therapeutic efficacy.[4,5] Therefore, developing nucleoside hydrogels with antibacterial properties and low biological toxicity using a systematic screening strategy is crucial for improving the therapeutic outcomes of periodontitis, and could also validate the results of our predictive models.

Based on predictive results (Fig. 2 a, Supplementary Data 8), Molecular Bioactivity Specificity Index (MBSI) was established to evaluate the dominant bioactivity of nucleoside derivatives (Supplementary Data 9), while the Composite Molecular Attribute Score (CMAS) was developed to assess the overall performance of antibacterial nucleoside hydrogels (Supplementary Data 10). The MBSI is used to evaluate the dominant or specific bioactivity within a molecule. One bioactivity with higher index within a molecule means that the particular bioactivity of the molecule is more prominent. This index could help identify which specific bioactivity within a molecule is most likely to excel, as well as pinpoint molecules that exhibit high activity in one bioactivity while remaining inactive in others. The CMAS is derived from the Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS),[26,27] and is designed for the comprehensive evaluation of molecules with multiple bioactivities. A higher score indicates that the molecule is closer to the ideal solution, representing better overall performance.

The screening strategy consisted of four steps (Fig. 3a). First, predictive models for hydrogel-forming ability, biotoxicity, and anti-*P. gingivalis* activity were constructed. Second, the MBSI and CMAS result for the three properties of nucleoside derivatives ($N = 6\,817$) were calculated (Supplementary Data 9, 10). The top 1% ($n = 68$) molecules with the highest CMAS scores were selected for further screening. Third, four candidate nucleoside derivatives were chosen based on chemical synthesizability (Guanosine - 5′ - monophosphate, GMP; 2′ - Deoxyguanosine 5′ - phosphate, dGMP; 2′ - O - Methylguanosine 5′ - monophosphate, 2′-OMe-dGMP; 5 - (1,3 - Dithiolan - 2 - yl) - 1 -((2R, 4S, 5R) - 4 - hydroxy - 5 - (hydroxymethyl) tetrahydrofuran - 2 - yl) pyrimidine - 2, 4(1H, 3H) - dione, 5′-Dithio-dGTP). After in vitro antibacterial performance test, GMP, dGMP, 2′-OMe dGMP, and 5′-Dithio-dGTP showed varying degrees of antibacterial activity, with GMP and

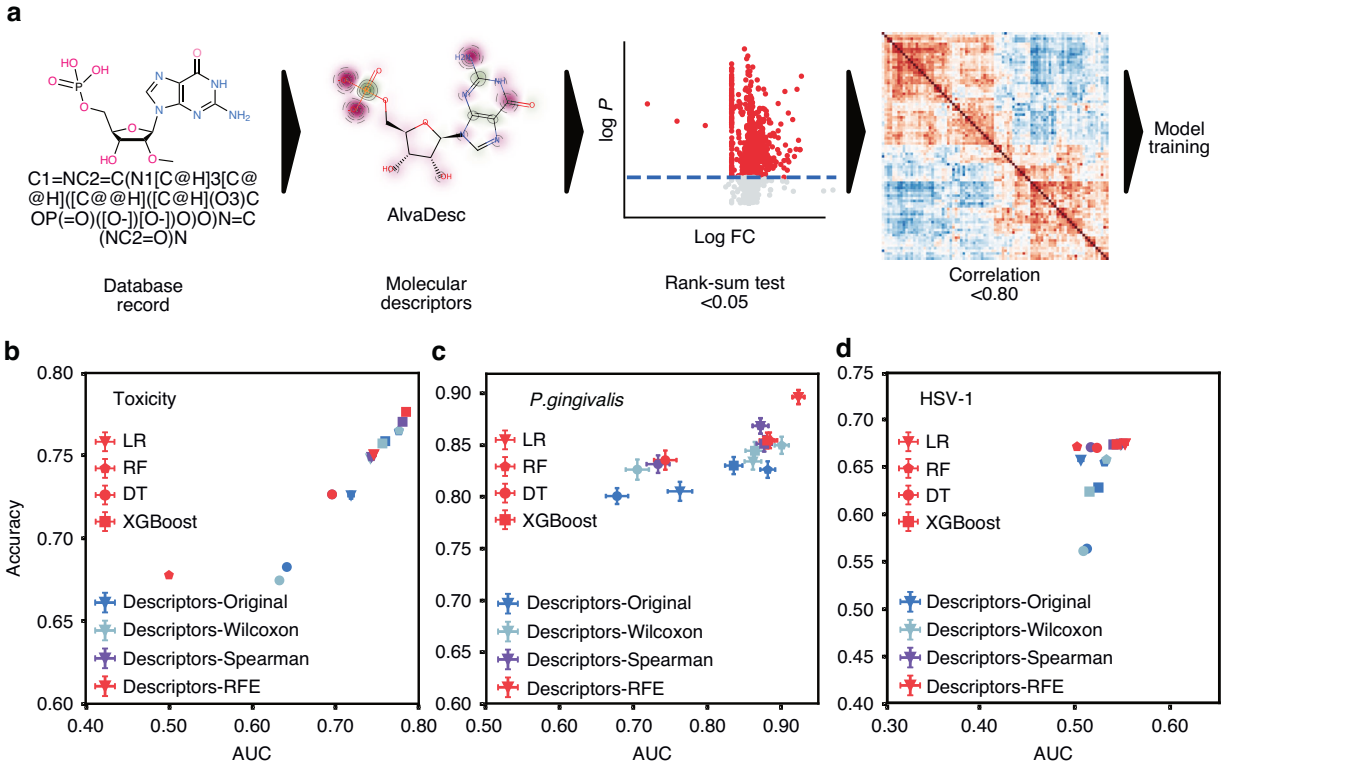

Fig. 1 Development of machine learning models. **a** Flowchart of feature selection in machine learning used for predicting nucleoside properties. **b-d** Scatter plots displaying the distribution of AUC (area under the curve) and test accuracy for Toxicity, *P. gingivalis* and HSV-1 models. The four point shapes represent different ML algorithms: extreme gradient boosting (XGBoost), logistic regression (LR), decision tree (DT), and random forest (RF). Descriptor parts include: initially obtained descriptors, descriptors after rank-sum test, descriptors after correlation coefficient selection, and descriptors after recursive feature elimination (RFE)

dGMP demonstrating superior efficacy (Fig. 3b, Supplementary Fig. 5, 6). Finally, gelation experiments were performed on the most promising nucleoside monomers (GMP, dGMP), followed by in vitro and in vivo validation of their bioactivity to determine the target nucleoside hydrogels (GMP hydrogel and dGMP hydrogel). For more details on the screening process, see Fig. 3a, b and Supplementary Data 8–10.

Hydrogel formation and mechanical properties of GMP and dGMP
To investigate the hydrogel-forming ability, strength, and stability of the two nucleotide derivatives (GMP and dGMP), we conducted a series of experiments including vial-inversion testing, scanning electron microscopy (SEM) imaging, Fourier-transform infrared (FTIR) spectroscopy, powder X-ray diffraction (PXRD), and rheological testing. The vial inversion test demonstrated the macroscopic gelation behavior, GMP and dGMP formed hydrogels in the presence of Ag^+ (Fig. 4a). SEM images revealed that both GMP and dGMP hydrogels possessed porous structures, with GMP hydrogel exhibiting a more ordered network compared to dGMP hydrogel (Fig. 4b). FTIR spectroscopy showed that gelation significantly altered the hydrogen bonding network between GMP and dGMP molecules, leading to reduced intensities of characteristic absorption peaks (Fig. 4c). PXRD analysis indicated that GMP and dGMP gels mainly presented an amorphous structure, with some diffraction peaks suggesting the presence of localized ordered regions formed through π-π stacking interactions between aromatic rings (Fig. 4d).[28] The self-healing ability of hydrogels is critical for biomedical applications. Rheological cyclic strain time sweep tests showed that GMP and dGMP hydrogels exhibited excellent self-healing properties, restoring their structure after strain-induced disruption (Fig. 4e). The mechanical properties of hydrogels were determined by rheological measurements. The elasticity and flowability properties were measured in terms of storage modulus and loss modulus, termed G′ and G″, respectively.[29] G′ was consistently G″ over a range of frequencies, indicating good mechanical strength and structural stability (Fig. 4f, g). Moreover, the hydrogels exhibited pronounced shear-thinning behavior, demonstrating good flow properties and processability under external forces (Fig. 4h).

Evaluation of antibacterial activity and cytocompatibility of GMP and dGMP Hydrogels against *P. gingivalis*
The antibacterial and biological toxicity activity of different concentrations of GMP and dGMP, as well as their hydrogel forms, were evaluated against *P. gingivalis*.[30,31] In Supplementary Fig. 5 **a**, increasing concentrations of GMP and dGMP show enhanced inhibitory effects on *P. gingivalis*, with significant bacterial growth inhibition observed at higher concentrations. Supplementary Fig. 5b shows the live/dead staining results of *P. gingivalis* under different treatments, where the number of live cells decreases while the proportion of dead cells increases as the concentration of GMP and dGMP increases, indicating improved antibacterial activity. Figure 5a, b further quantify the inhibitory effects of different concentrations of GMP and dGMP, demonstrating a significant reduction in bacterial colony count at concentrations of 1 mmol/L and above, leading to nearly complete inhibition of bacterial growth. In Fig. 5c SEM images illustrate the morphology of *P. gingivalis* treated with PBS, Ag^+, GMP and dGMP hydrogel. The bacteria in the PBS group maintained an intact morphology, whereas those treated with Ag^+, GMP hydrogel, and dGMP hydrogel showed significant surface damage, indicating strong bactericidal effects of these treatments. Figure 5d demonstrates the effect of Ag^+, GMP hydrogel, and dGMP hydrogel on bacterial colony counts, showing a significant

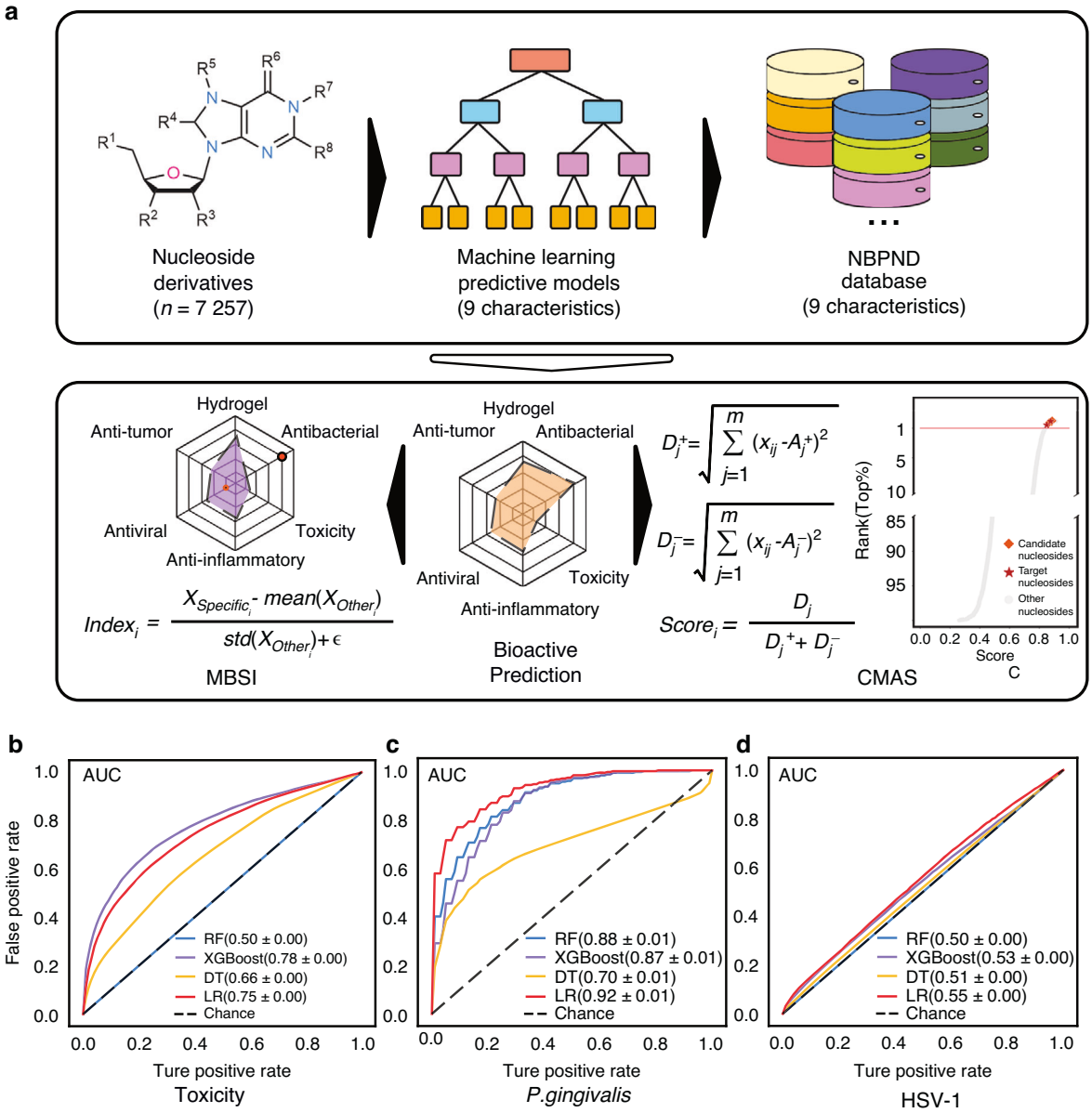

Fig. 2 Workflow and predictive performance of nucleoside derivatives. **a** Flowchart depicting the prediction of nucleoside derivatives with specific biological activity and physical properties. NBPRP New biological and physical properties predicted databases, MBSI molecular bioactivity specificity index, CMAS composite molecular attribute score. **b–d** AUC (Area under the curve) results for four algorithms (LR, DT, RF, and XGBoost) using descriptors after RFE

reduction in bacterial numbers compared to the PBS control, confirming their effective antibacterial activity. Figure 5e shows the results of cell viability tests, where Ag^+ treatment significantly reduced cell viability, indicating biological toxicity, while GMP and dGMP hydrogels maintained nearly 100% cell viability, demonstrating excellent biocompatibility. These results suggest that GMP and dGMP exhibit strong antibacterial effects, especially in their hydrogel forms, which show good antibacterial activity and cytocompatibility. *P. gingivalis* is a keystone periodontal pathogen, and its overgrowth is central to the initiation and progression of periodontitis. Therefore, the strong antibacterial activity demonstrated by GMP and dGMP is directly linked to their therapeutic relevance. By effectively inhibiting *P. gingivalis*, these hydrogels help restore microbial balance in the periodontal pocket, which is essential for suppressing inflammation and preventing further periodontal tissue destruction.

In vivo biocompatibility and organ toxicity of GMP and dGMP gels The in vivo biocompatibility test results indicate that inflammation at the injection sites of GMP and dGMP gels gradually returned to normal within 7 days post-injection (Supplementary Fig. 7). Hematoxylin and eosin (H&E) staining results showed intact skin tissue structures without significant pathological changes at the injection sites, demonstrating good in vivo biocompatibility of both hydrogels. These findings suggest that GMP and dGMP hydrogels exhibit excellent biocompatibility following subcutaneous injection, without causing noticeable tissue damage or persistent inflammation. The biocompatibility evaluation in major organs results show that in all treatment groups, including GMP, and dGMP, as well as their hydrogels, no inflammatory infiltration was observed in major organs such as the heart, spleen, liver, lung, and kidney (Supplementary Fig. 8). All organs maintained an intact overall structure without any significant pathological changes or

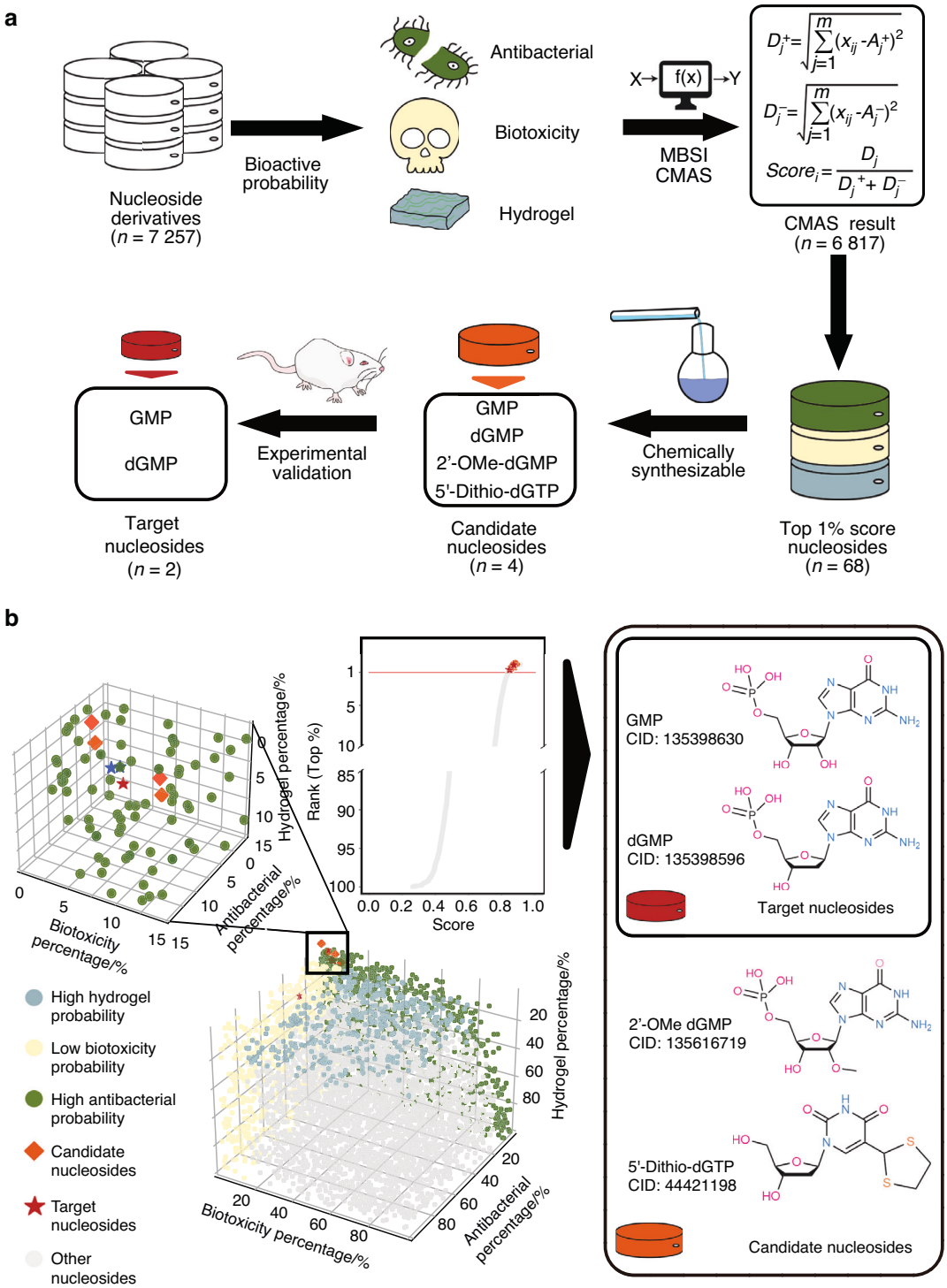

Fig. 3 Screening workflow and candidate selection of nucleoside derivatives for periodontitis treatment. **a** Screening of nucleoside hydrogels for periodontitis treatment based on hydrogel-forming ability, biotoxicity (biological toxicity), and antimicrobial properties. MBSI molecular bioactivity specificity index; CMAS composite molecular attribute score. **b** The screening process was based on predicted probability of hydrogel-forming ability, biotoxicity, and the anti-*P. gingivalis* activity. Rankings were determined using the CMAS method, and four candidate nucleoside derivatives were selected from the top 1% of nucleoside derivatives ($n = 68$) based on their chemical synthesizability

tissue damage, indicating that the treatments did not induce inflammation or organ toxicity. These results indicate that GMP and dGMP hydrogels have good biocompatibility in major organs of mice, without causing noticeable tissue damage or inflammatory response, demonstrating their safety for further biomedical applications. The demonstrated in vivo biocompatibility and absence of organ toxicity further underscore their suitability for periodontal use, as materials designed to remain within periodontal pockets must ensure both local tissue safety and systemic tolerance.

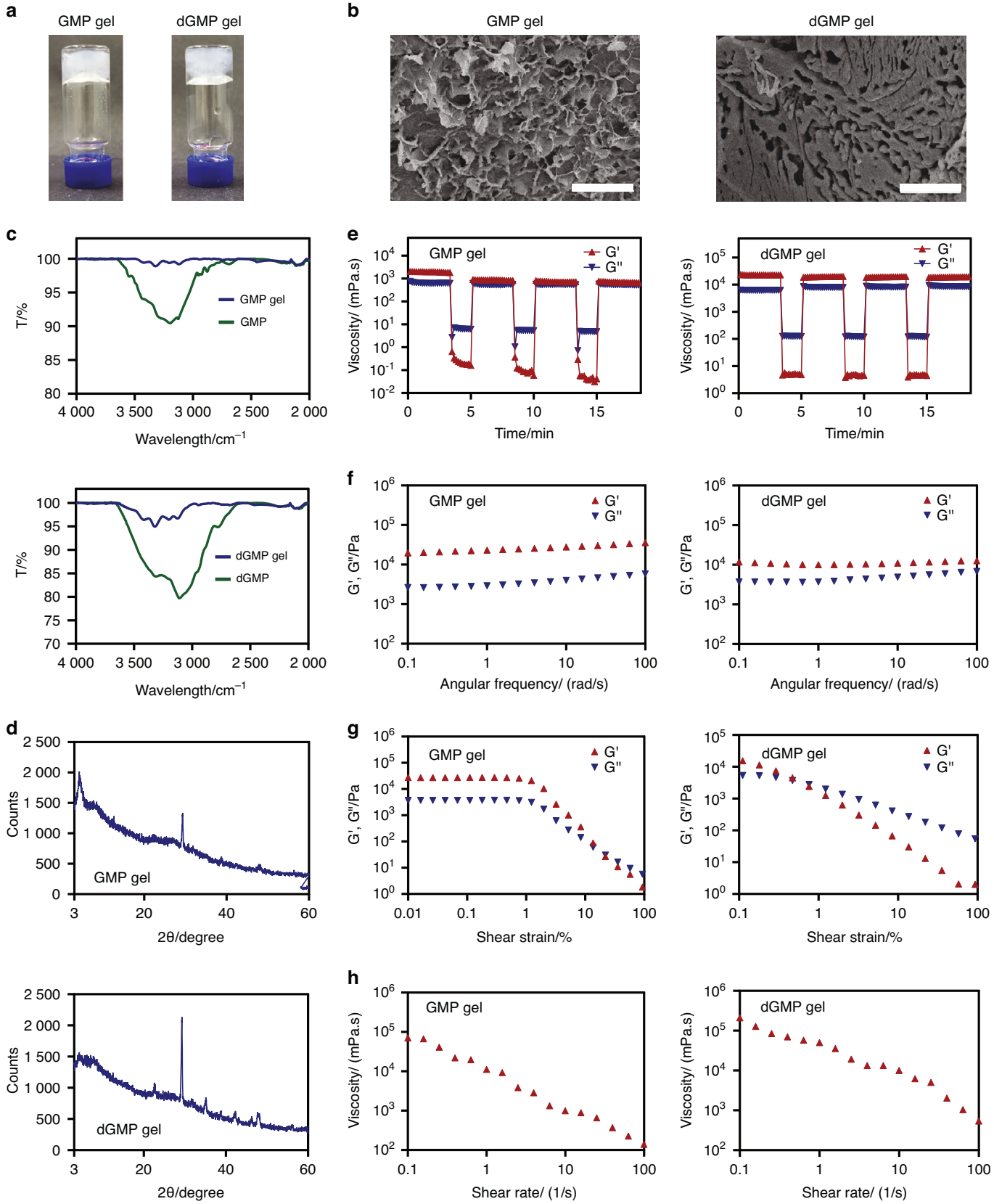

Fig. 4 Physical and rheological properties of GMP and dGMP hydrogels. **a** Vial inversion test demonstrating the gel-forming ability of GMP and dGMP gels. **b** Scanning electron microscopy (SEM) images showing the porous structures of GMP hydrogel, and dGMP hydrogel. **c** Infrared spectroscopy (IR) spectra comparing the chemical bonds present in GMP hydrogel, and dGMP hydrogel. **d** X-ray diffraction (XRD) analysis depicting the crystallinity of GMP hydrogel and dGMP hydrogel. **e-h** Rheological properties of GMP hydrogel, and dGMP hydrogel: **e** Self-healing ability as assessed by repeated strain tests; **f** Frequency sweep showing the storage modulus (G′) and loss modulus (G″); **g** Strain sweep test indicating the mechanical stability; **h** Shear thinning behavior observed through viscosity changes under varying shear rates

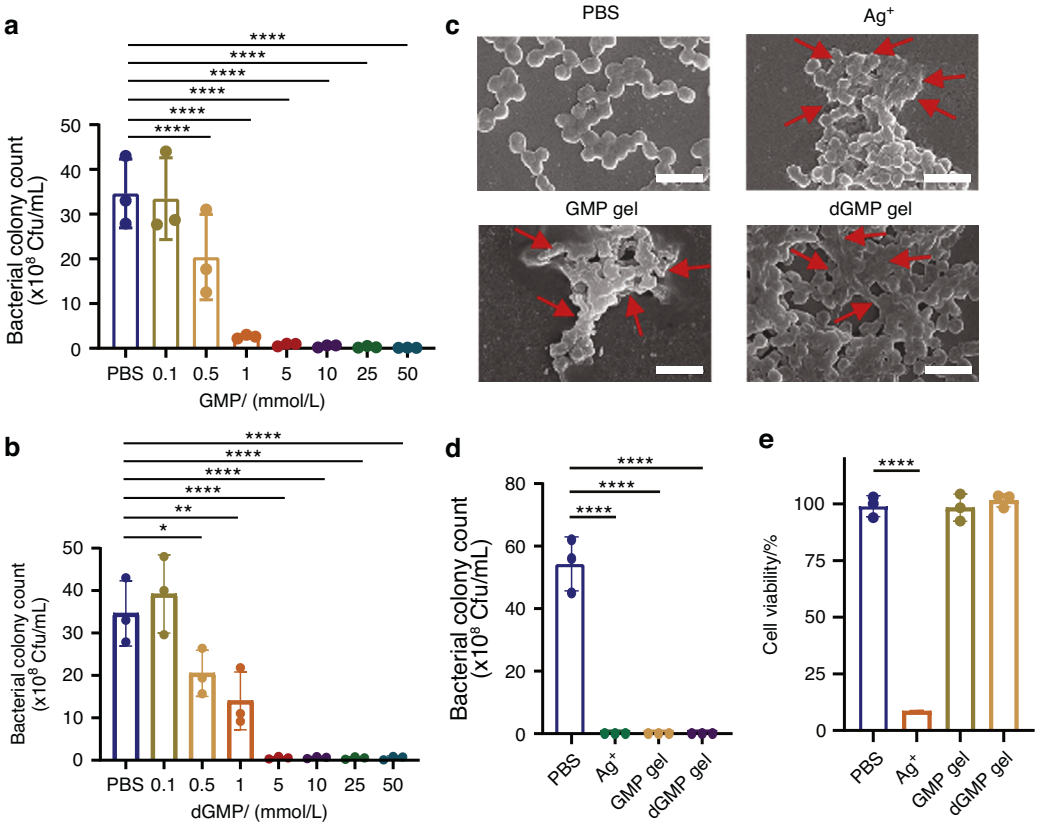

Fig. 5 In vitro evaluation of antibacterial activity and cytotoxicity of GMP and dGMP hydrogels. **a** Quantitative analysis of CFU counts for GMP monomers at different concentrations. **b** Quantitative analysis of CFU counts for dGMP monomers at different concentrations. **c** SEM images of tooth samples treated with different gel formulations: PBS, Ag^+, GMP hydrogel, and dGMP hydrogel. **d** Colony count comparison among PBS, Ag^+, GMP hydrogel, and dGMP hydrogel groups. **e** Cell viability of fibroblast cells after treatment with PBS, Ag^+, GMP hydrogel, and dGMP hydrogel

Periodontitis treatment model based on inhibition of *P. gingivalis*
A mouse model of periodontitis was established to evaluate the therapeutic effects of GMP and dGMP hydrogels on periodontitis through the inhibition of *P. gingivalis*. After one week of acclimation, mice were subjected to 5-0 silk ligature with *P. gingivalis* to induce periodontitis. Fourteen days post-infection, different treatments, including 100 µL PBS, Ag^+, GMP monomers and dGMP monomers, GMP hydrogels and dGMP hydrogels (at concentrations of 50 and 100 mM), were injected into the periodontal pockets. The mice were then observed for an additional 14 days, followed by evaluations using Micro-CT analysis, H&E staining, and immunohistochemistry (IHC, including TNF-α and TGF-β) to assess the inhibition of *P. gingivalis* (Fig. 6a).

Micro-CT images showed that alveolar bone loss was significantly reduced in the hydrogel treatment groups compared to the periodontitis model group (Fig. 6b).[24] Quantitative parameters, including alveolar bone crest to cemento-enamel junction (ABC-CEJ), Bone Volume to Total Volume (BV/TV), trabecular separation (Tb.Sp), trabecular thickness (Tb.Th), and trabecular number (Tb.N), further supported this conclusion (Fig. 6c-d, Supplementary Fig. 9). These results indicating that GMP and dGMP hydrogels effectively mitigated bone loss and increased trabecular thickness and number, with more pronounced effects observed in the high-concentration hydrogel groups. H&E staining (Fig. 7a, Supplementary Fig. 10) indicated that the inflammatory response was reduced in the hydrogel treatment groups compared to the periodontitis model group. IHC results for TNF-α and TGF-β (Fig. 7a, Supplementary Fig. 11-12) further demonstrated that TNF-α expression, an indicator of inflammation, was significantly lower in the hydrogel groups, while TGF-β, a marker of tissue repair and regeneration, was increased, especially in the high-concentration hydrogel groups. Quantitative analysis (Fig. 7b-c, Supplementary Fig. 13) showed that the AOD (average optical density) and percentage of positive staining for TNF-α were significantly reduced in the GMP and dGMP hydrogel groups, whereas TGF-β expression was notably increased, indicating effective suppression of inflammation and promotion of tissue healing.[24]

The observed recovery of alveolar bone structure and reduction in inflammation markers could demonstrate the therapeutic efficacy of GMP and dGMP hydrogels in treating established periodontitis, aligning the characterization results with meaningful clinical periodontal outcomes.

Periodontitis prevention model based on inhibition of *P. gingivalis*
Figure 8a illustrates the experimental process for the periodontitis mouse prevention (or pre-treatment) model. Initially, mice were acclimated for seven days, followed by induction of periodontitis on day 0 through ligature placement and topical application of *P. gingivalis*. On the next day, hydrogels were injected into the periodontal pockets to evaluate their preventive effects against *P. gingivalis*-induced periodontitis.

The Micro-CT images indicate that GMP and dGMP hydrogels, similar to the positive control, effectively preserved alveolar bone compared to the untreated periodontitis group (Fig. 8b). Figure 8c, d and Supplementary Fig. 14 illustrate the effects of GMP and dGMP hydrogels (100 mM) on alveolar bone preservation in a periodontitis model, with minocycline serving as a positive control. In the hydrogel-treated groups, the ABC-CEJ distance (Fig. 8c) was reduced, and bone volume fraction (BV/TV, Fig. 8d) increased significantly. Additionally, trabecular separation,thickness and

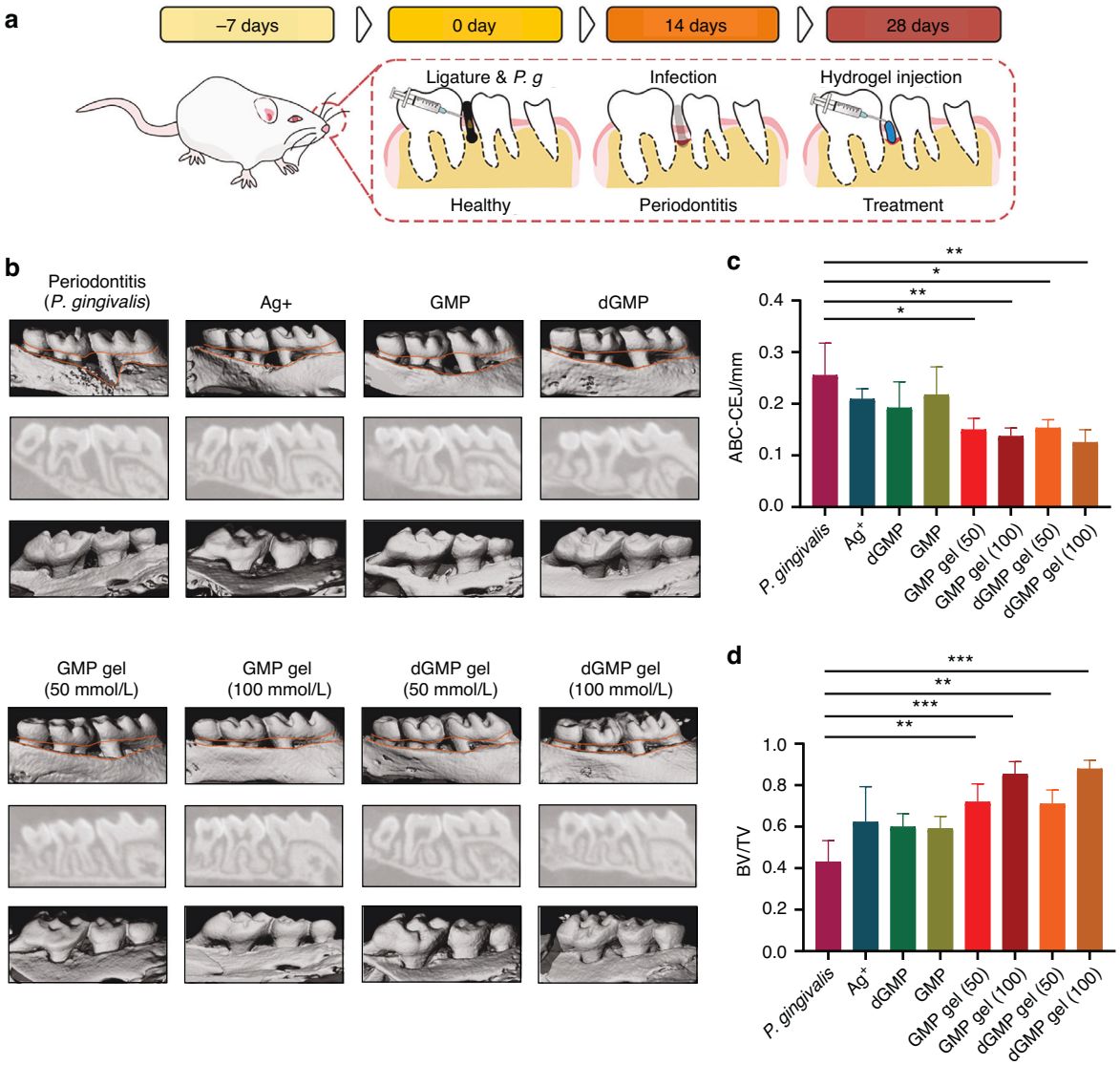

Fig. 6 Evaluation of of nucleoside hydrogels in a periodontitis treatment model. **a** Experimental scheme showing the timeline of ligature-induced periodontitis, treatment with *P. gingivalis*, and subsequent hydrogel treatment. **b** Micro-CT images demonstrating alveolar bone loss in different treatment groups, including *P. gingivalis*, Ag^{+}, GMP monomer, dGMP monomer, GMP hydrogel, and dGMP hydrogels. **c–d** Quantitative analyses of alveolar bone crest to cemento-enamel junction (ABC-CEJ) distance, and bone volume fraction (BV/TV)

trabecular number (Tb.Sp, Tb.Th, Tb.N, Supplementary Fig. 14a-c) were both enhanced. Overall, GMP and dGMP hydrogels demonstrated effective prevention of bone loss and improved bone quality, comparable to the effects of the positive control minocycline. H&E staining showed a marked reduction in inflammation in the hydrogel-treated groups compared to the *P. gingivalis*-induced periodontitis group, with tissue structure resembling that of the healthy controls (Fig. 9a). These observations were consistent with the positive control group. IHC analysis for TNF-α and TGF-β further demonstrated the benefits of the GMP and dGMP hydrogels (Fig. 9b-e, Supplementary Fig. 15). TNF-α, an indicator of inflammation, was significantly reduced, while TGF-β, associated with tissue healing, was notably increased in the GMP and dGMP hydrogel groups. The results suggest that GMP and dGMP hydrogels effectively reduced the periodontitis induced by *P. gingivalis*.

These findings highlight that early administration of GMP and dGMP hydrogels prevents alveolar bone loss and suppresses *P. gingivalis*-induced inflammation. Furthermore, the comparable performance of these hydrogels to minocycline demonstrates their strong prophylactic potential and suggests they may serve as promising alternatives or adjuncts to clinically used periodontal therapeutics.

DISCUSSION

This study successfully developed ML-based predictive models to forecast nine biological properties of nucleosides, enabling the design and validation of nucleoside hydrogels for specific bioactivity applications. Through a comprehensive screening based on CMAS and an experimental validation process, two novel nucleoside hydrogels were identified and shown to be effective for the treatment of periodontitis. These findings not only provide insight into the structure-property relationships of nucleosides but also demonstrate the practical application of the synthesis strategy based on predictive models and MBSI/CMAS methods in bioactive supramolecular nucleoside hydrogels development.

The biological properties of nucleoside hydrogels, such as antibacterial activity, and biocompatibility, directly determine their

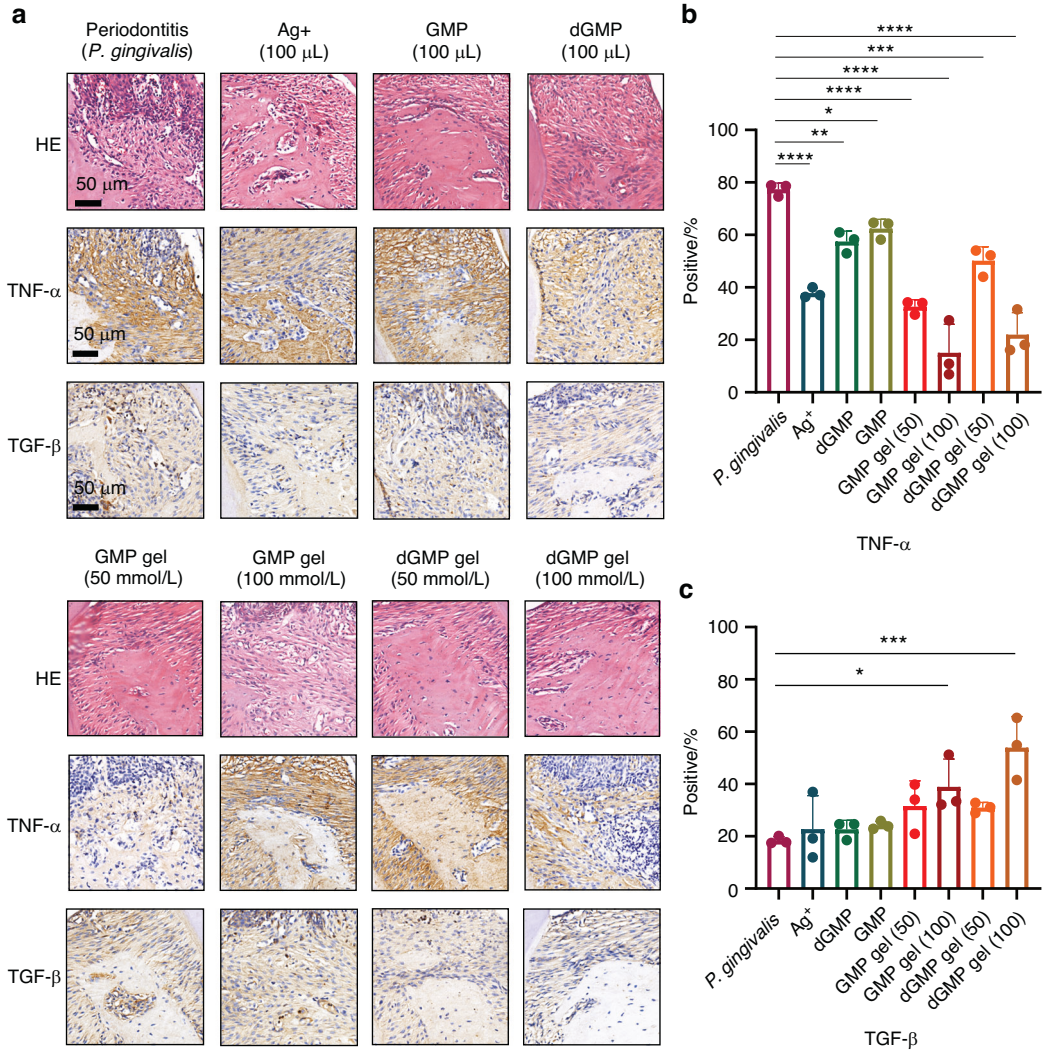

Fig. 7 Evaluation of gingival inflammation and tissue healing in a periodontitis treatment model. **a** Histological examination with H&E staining of gingival tissue showing inflammation levels after treatment with different gels, Immunohistochemical (IHC) analysis of TNF-α and TGF-β expression in gingival tissue, indicating inflammation status across various groups. **b–c** IHC quantitative assessments of TNF-α and TGF-β expression

efficacy in different applications. For example, hydrogels for drug delivery systems require excellent gelation performance to ensure controlled and stable drug release at target sites,[3–5,7] while in wound healing or inflammation treatment, strong antibacterial activity and low biological toxicity are crucial.[4–6] The structural complexity and diversity of nucleosides make it challenging to develop functional hydrogels using traditional synthesis and screening methods. Which often rely on extensive trial-and-error experimentation, consuming significant time and resources. Moreover, the relationship between nucleoside structure and physicochemical properties remains poorly understood, further complicating the design of hydrogels with specific functionalities.[8–10] Therefore, a systematic synthesis strategy is urgently needed to guide the development of nucleoside hydrogels with tailored properties for precise biomedical applications.[10]

The synthesis strategy based on ML models and MBSI/CMAS methods developed in this study allowed us to efficiently select nucleoside derivatives with optimal gelation, antibacterial, and biocompatibility properties.[8,10,11] Compared to traditional extensive trial-and-error experiments, this synthesis strategy provided a systematic, data-driven framework to identify promising candidates. By integrating large-scale datasets, the models captured complex relationships between nucleoside structure and their properties, MBSI/CMAS methods improved the efficiency and precision of the screening process.[10–13,19] This approach can be extended beyond periodontitis treatment, offering a powerful tool for the rational design of nucleoside hydrogels for various therapeutic applications. The predictive database developed in this study aggregates physicochemical and biological data for a wide range of nucleosides, providing researchers with a valuable resource that enables rapid prediction of nucleoside properties without the need for exhaustive experimentation.

In this study, MBSI and CMAS are proposed for the first time to assess the bioactivities of nucleoside molecules, providing a systematic and scientific approach for screening ideal bioactive nucleoside hydrogels. These methods establish a comprehensive framework for guiding the design and development of nucleoside hydrogels with specific biological functions. The MBSI evaluates the dominant or specific attributes of molecules, enabling the identification of those that excel in specific bioactivities or are active in one function while inactive in others. The CMAS, based on the TOPSIS, allows for a comprehensive evaluation of molecules with multiple properties. Based on this, a systematic screening process was established to assess multiple bioactivities. This approach allows for the rapid identification of nucleoside

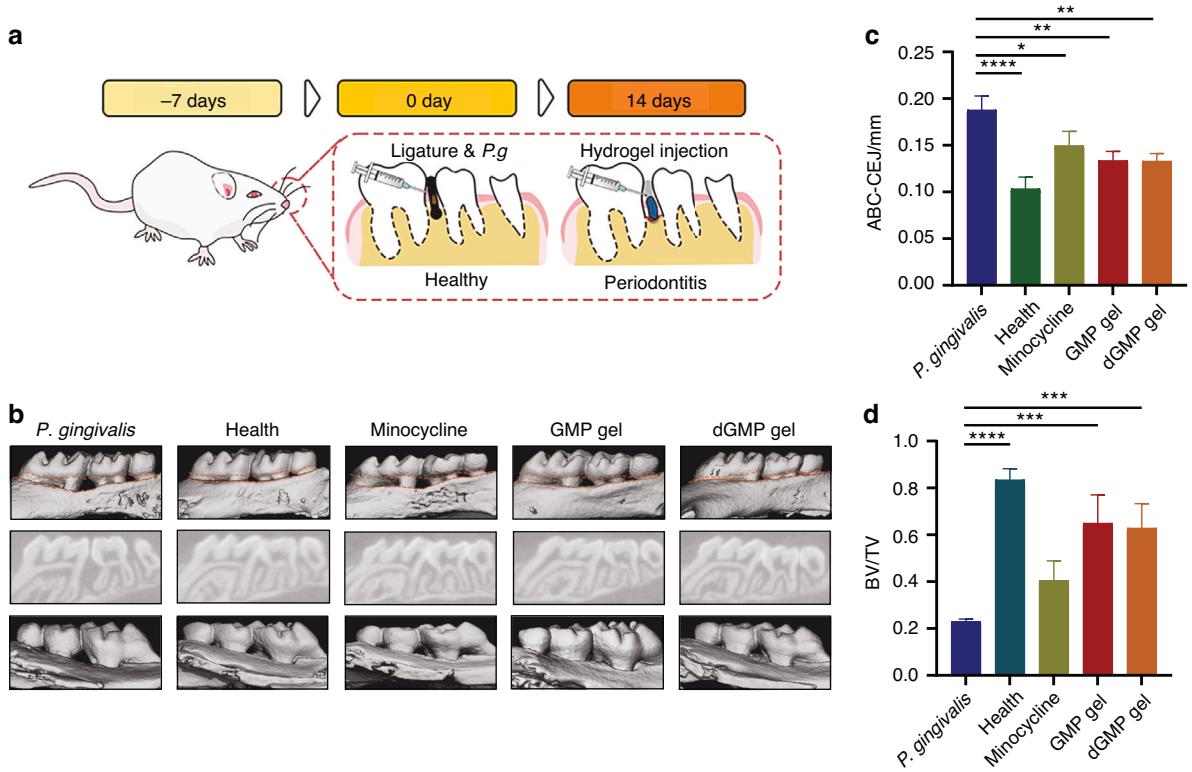

Fig. 8 Evaluation of of nucleoside hydrogels in a periodontitis prevention model. **a** Schematic of ligature-induced periodontitis and subsequent hydrogel application, showing different stages from ligature to treatment. **b** Micro-CT images assessing the extent of periodontal bone loss across experimental groups, including *P. gingivalis*, minocycline (positive control), GMP hydrogel and dGMP hydrogel. **c**, **d** Quantitative assessments of ABC-CEJ distance, BV/TV

hydrogels suitable for specific clinical applications, thus achieving the goal of scientifically synthesizing nucleoside hydrogels with targeted bioactivities. Using ML and other AI algorithms for multi-property predictive analysis has become a major research focus. However, the integration of multiple properties to predict and select the optimal multi-property molecule has not been widely reported. This study makes the first attempt to advance this area by employing MBSI and CMAS. The synthesis strategy based on ML models and MBSI/CMAS methods is not limited to nucleoside hydrogels with specific bioactivities; it is also applicable to the prediction and development of a broad range of molecules with targeted bioactivities. Notably, for periodontal applications, materials must exhibit antibacterial potency, low toxicity, and strong tissue-compatibility simultaneously, making this multi-dimensional assessment particularly valuable.

Although Ag^+ was included in the gelation system, our current evidence does not confirm whether it plays a decisive role in the assembly of GMP or dGMP hydrogels. The hydrogels exhibited markedly stronger antibacterial and therapeutic effects than either Ag^+ alone or the corresponding nucleoside monomers, suggesting that the supramolecular network rather than Ag^+ itself contributes predominantly to the observed bioactivity, with the possibility of synergistic effects that warrant further investigation. Importantly, histological analyses of major organs and injection sites revealed no tissue damage, indicating that the residual Ag^+ did not induce measurable cytotoxic effects in our experimental setting.

The experimental validation confirmed the high performance of the selected nucleoside hydrogels in periodontitis treatment. GMP and dGMP hydrogels demonstrated excellent gelation ability, forming stable networks with high mechanical strength, which is crucial for effective drug delivery and retention in periodontal tissues. Additionally, the hydrogels exhibited strong antibacterial activity, effectively inhibiting the growth of *P. gingivalis*, a major pathogen associated with periodontitis. The biocompatibility tests further supported their suitability for biomedical use, as both hydrogels showed minimal biological toxicity and high cell viability, ensuring their safety in clinical applications. The successful in vivo application of the hydrogels in animal models underscores their therapeutic potential. Both hydrogels significantly reduced bacterial load, suppressed inflammation, and promoted tissue regeneration, demonstrating their ability to address the *P. gingivalis*-induced periodontitis. These results suggest that nucleoside hydrogels can provide a multi-faceted approach to treatment by combining antibacterial effects, which is crucial for managing chronic inflammatory diseases like periodontitis. These characterizations, particularly the improvements in alveolar bone morphology (ABC–CEJ, BV/TV, Tb.Th, Tb.N), reductions in inflammatory markers (TNF-α), and increases in regenerative markers (TGF-β)—provide quantitative evidence of therapeutic efficacy, allowing the treatment effect to be inferred from measurable biological endpoints. Additionally, the study highlights the potential of nucleoside-based biomaterials to improve outcomes in other areas of oral healthcare, such as peri-implantitis or postoperative wound management.

Although the results are promising, there are some limitations to this study. Firstly, the biological activity predictions were based on existing known information of common compounds, lacking a dedicated dataset of nucleoside derivative bioactivities for model training. Secondly, due to experimental limitations, it was challenging to perform extensive external validation of the antibacterial, toxicity, and other bioactivities of the nucleoside hydrogels. Although in vivo studies provided valuable insights into the therapeutic potential of the hydrogels, further clinical trials are needed to validate their efficacy and safety in human populations. Finally, the long-term stability and biodegradability of the hydrogels

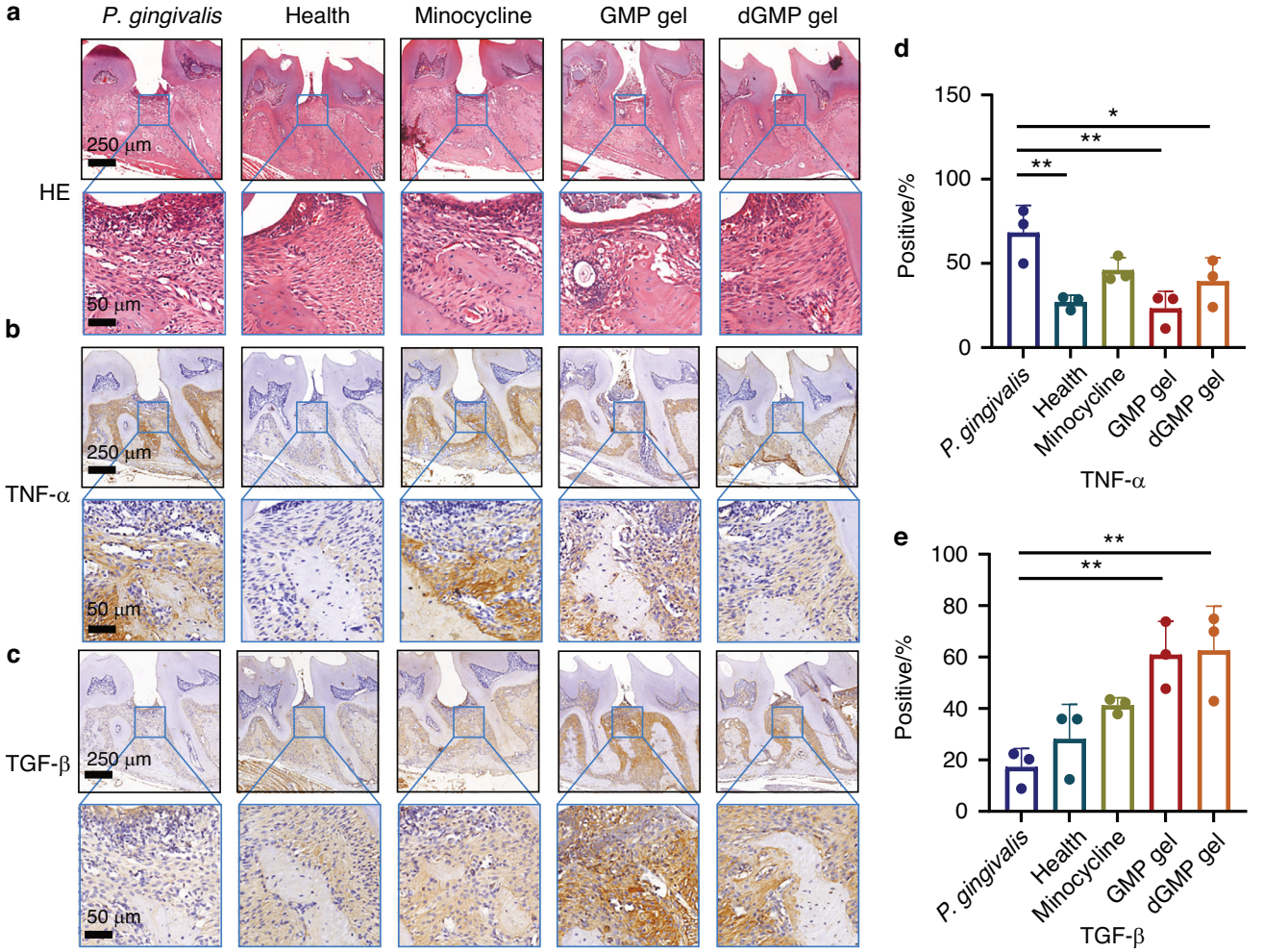

Fig. 9 Evaluation of gingival inflammation and tissue healing in a periodontitis prevention model. **a** Hematoxylin and eosin (H&E) staining of the gingival tissue. **b** Immunohistochemical (IHC) analysis of TNF-α expression to determine inflammatory response in gingival tissues. **c** IHC analysis of TGF-β levels, evaluating tissue regeneration and anti inflammatory effects across groups. **d**, **e** Quantitative assessments of positive staining percentage for TNF-α and TGF-β

in complex biological environments require further research to ensure their suitability for prolonged clinical use.

This study represents only an initial attempt in the design and development of biological functional nucleoside hydrogels. Future research could expand the database by incorporating more nucleoside derivatives, while also improving predictive algorithms through techniques such as transfer learning or deep learning to enhance model performance. Moreover, the database could be integrated with chemical synthesis tools to provide end-to-end solutions for hydrogel development. Future research could explore several avenues to enhance the application of nucleoside hydrogels. Expanding the scope of predictive models to include properties like controlled release kinetics or immune modulation could open new possibilities for hydrogel design. Finally, developing personalized hydrogels tailored to individual patient needs through a synthesis strategy based on ML models and MBSI/CMAS methods could significantly advance the field of precision medicine.

In conclusion, this study demonstrates the effectiveness of a synthesis strategy based on ML models and MBSI/CMAS methods for in the design and development of bioactive nucleoside hydrogels. The combination of predictive modeling, database development, and experimental validation provides a robust framework for the rational synthesis of biomaterials with targeted properties. The successful application of these hydrogels in ther treatment of periodontitis underscores their potential for broader use in biomedical applications. By linking quantitative imaging, histological, and molecular characterization with periodontal outcomes, this study provides a study basis for inferring the therapeutic effects of nucleoside hydrogels on periodontitis. By leveraging data-driven approaches, this research paves the way for the future of hydrogel-based therapies and offers valuable insights into the rapidly field of biomaterials.

METHOD

Model construction

Data collection and preprocessing. We collected compound bioactivity data from databases such as PubChem and DeepChem to train ML models, ultimately selecting nine databases for model construction. DeepChem aims to provide a high-quality open-source toolchain to democratize the use of deep learning in drug discovery, materials science, quantum chemistry, and biology.[16] We used the biological toxicology dataset from Tox21 provided by DeepChem.[17] PubChem, managed by the National Institutes of Health (NIH), is a public repository for small-molecule biological properties.[15] We collected bioactivity datasets related to anti-bacterial (inhibitors of *P. gingivalis*), anti-viral (inhibitors of HSV-1, VZV, EBV, and HIV), anti-inflammatory (inhibitors of IL-1β and TNF-α), and anti-tumor activities (inhibitors of oral cancer cell line CAL 27). Categorical variables were converted into

numerical representations through one-hot encoding. Detailed information regarding the collection and preprocessing of the nine datasets can be found in Supplementary Methods and Supplementary Data 1.

Calculation of descriptors for bioactivity datasets. Both DeepChem and PubChem provide the SMILES (Simplified Molecular Input Line Entry System) strings for compounds (Supplementary Data 2). Molecular descriptors were calculated based on the Canonical SMILES of the compounds. In this study, a total of 5666 molecular descriptors were computed for compounds from the nine bioactivity datasets using the Python package alvaDescCLIWrapper (version 1.1.1) with Python version 3.9.12 (Supplementary Data 3).[10]

Feature selection based on descriptors. Referring to the previous studies, the three-step of feature selection method has been widely used in molecular property prediction models. Based on prior research experience, descriptors with more than 80% missing values were excluded,[10] and feature selection was carried out in three steps to prevent overfitting and improve model accuracy.[18,19] First, the rank-sum test was used as a univariate analysis to assess descriptor differences. Molecular descriptors that did not show significant association with the corresponding bioactivity ($P < 0.05$) were removed. Second, the Spearman correlation coefficient was calculated among the descriptors that passed the first step. Descriptor pairs with a correlation coefficient greater than 0.8 (Rho > 0.8) were identified, and one descriptor from each highly correlated pair was excluded to avoid multicollinearity. Finally, RFE was applied to obtain the optimal combination of remaining descriptors to maximize model performance. After 3-step feature selection, the number of features of the models have been reduced to an acceptable range. Hyperparameter optimization was conducted using the Bayesian optimization approach implemented in the Python package Optuna (version 3.1.0).[32]

In addition, we compared the three-step feature selection pipeline with a LASSO-based strategy, which was widely used in feature selection field.[33] For the LASSO approach, descriptors with zero values in more than 99% of molecules were first removed, and LASSO regression was then used to select informative descriptors before model training. Across the nine bioactivity tasks, LASSO-based models did not show superior test performance compared with the three-step method, which generally achieved higher test accuracy and AUC and smaller train–test resolution discrepancy (Supplementary Data 5). Feature importance of descriptors after feature selection by LASSO could be found in Supplementary Data 6.

Model training and evaluation. DT,[34] LR,[35] RF,[36] and XGBoost[37] models were implemented in Python using Scikit-learn (version 1.1.1).[38] RFE based on four ML algorithms was performed using a 10-times 5-fold stratified cross-validation scheme. After recursive feature elimination, the number of features will smaller than the sample size and reducing the risk of severe overfitting. Hyperparameter optimization was conducted based on the training set, using fivefold stratified cross-validation to estimate the 95% confidence interval for five evaluation metrics.[10] The data were divided into five equal parts while maintaining the ratio between active and inactive groups, with four parts used as the training set and the remaining part as the test set, repeating this process five times.[39] Additionally, we performed 10 iterations of the cross-validation to estimate model parameters, ensuring the stability of model performance and the reliability of results.[18,19] During training and evaluation, each model was trained on the training set and then evaluated on the test set to determine its ability to predict unseen data. Therefore, the training and testing scores for each model represent the average of these 10 iterations of five-fold cross-validation. A probability threshold of 50% was used for maximum likelihood estimation during both training and testing phases.

Determination of the optimal predictive model. The datasets obtained from the four stages of feature selection were used to train four different machine learning models. The goal was to identify the optimal predictive model for each of the nine bioactivities. The selection of the optimal model primarily focused on the test accuracy and AUC of each model, while also considering precision, recall, and F1 score. These details are provided in Supplementary Method 2 and Supplementary Table 1.

Screening strategy. The screening strategy involved four steps. Firstly, predictive models for biotoxicity, gelation performance, and anti-P.gingivalis activity were constructed. Subsequently, the result of MBSI and CMAS for the three properties of nucleoside derivatives was calculated. The top 1% molecules with the highest CMAS scores were selected for further screening. Thirdly, candidate nucleoside derivatives with high chemical synthesizability were selected from the top 1% of molecules with the highest CMAS scores for synthesis. Finally, the most promising nucleoside monomers were subjected to gelation experiments, and their bioactivity was validated in vitro and in vivo to identify the target nucleoside hydrogel.

Molecular bioactive performance analysis

Molecular bioactivity specificity index. This method uses Z-score standardized bioactivity data to calculate the distance between a specific attribute and the overall distribution of other attributes, which is considered as the specificity of the bioactivity. It effectively reveals the most prominent molecular feature in each molecule and provides a comprehensive specificity score, enhancing our understanding of the differences and underlying biological mechanisms of molecular bioactivities. Specifically, the MBSI represents the dominant bioactivity of a molecule. In this study, it was used to evaluate the specificity of nucleoside derivatives in terms of biocompatibility and anti-*P. gingivalis* activity. A higher index indicates higher specificity in the given bioactivity. Beyond evaluating the dominant bioactivities of nucleoside derivatives discussed in this study, this method can also be applied to access the specificity of molecules across multiple attributes, providing valuable guidance for the rational synthesis of molecules with targeted properties.

Standardizing the data matrix. First, we standardize the data matrix using Z-score normalization, where each attribute value is standardized by subtracting the mean and dividing by the standard deviation. To avoid division by zero, a small constant ϵ is added to the standard deviation:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j - \epsilon} \quad (1)$$

Where:

z_{ij} is the value of molecule i in bioactivity j of the original data matrix.

μ_j is the mean of bioactivity j.

σ_j is the standard deviation of bioactivity j.

Identifying the most specific feature. For each sample, the feature with the highest standardized value is identified as the most specific bioactivity :

$$Specific_i = \arg\max_j (z_{ij}), i = 1, 2, \ldots, m \quad (2)$$

$Specific_i$ is the most specific bioactivity of molecule.

m is the number of bioactivities.

Calculating the specificity index. To quantify the prominence of the most specific bioactivity, the MASI Feature Index is calculated. This involves comparing the value of the most specific bioactivity

to the mean and standard deviation of the other bioactivities in the same molecule:

$$Index_i = \frac{X_{Specific_i} - mean(X_{Other_i})}{std(X_{Other_i}) + \epsilon} \quad (3)$$

$X_{Specific_i}$ is the value of the most specific bioactivity of molecule *i*.

X_{Other_i} represents the values of all other bioactivities in the same sample of molecule *i*.

In summary, this method identifies the dominant bioactivity of molecules, quantifies their specificity, and assists in evaluating and guiding the targeted synthesis of bioactive nucleoside derivatives.

Composite molecular attribute score

This method utilizes min-max normalized bioactivity data to evaluate the proximity of each attribute to ideal and negative-ideal solutions, providing a comprehensive CMAS score that reflects the relative performance of molecular features. Specifically, the CMAS score quantifies how closely a given feature aligns with the optimal bioactivity. In this study, the CMAS analysis was applied to assess the bioactivity of nucleoside derivatives, focusing on biocompatibility and anti-*P. gingivalis* activity. A higher CMAS score indicates better alignment with the desired bioactivity. Beyond the bioactivities of nucleoside derivatives examined here, CMAS can also be applied to evaluate molecular attributes across diverse features, offering valuable insight for designing molecules with desired bioactive properties.

Normalization of weighted matrix. First, we normalize the data matrix using Min-Max normalization, followed by applying a specific weight to each feature.

The normalized and weighted data:

$$x_{ij} = \omega_j \cdot \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}, i = 1, 2, \ldots, n; j = 1, 2, \ldots, m \quad (4)$$

Row *i* represents the molecule with bioactive predicted results, and column *j* represents the predicted bioactivity. The default and recommended number of properties is $m = 3$ (representing gelation ability, cytotoxicity, and selected bioactivity). However, any number of bioactivities ($m \geq 2$) can be set, as long as they correspond to the predicted properties provided.

x_{ij} is the value of molecule *i* in bioactivity *j* of the original data matrix.

n is the number of samples, *m* is the number of features.

ω_j represents the weight for bioactivity *j*, by default, it is assumed that the weights of the *m* features are equal (each being $1/m$). However, independent weights can be assigned to each bioactivity, provided that their sum equals 1.

Calculating ideal and negative-ideal solutions and the distance. Ideal solution (A_j^+)

$$A_j^+ = \max(x_{ij}), j = 1, 2, \ldots, m \quad (5)$$

Negative-ideal solution (A_j^-):

$$A_j^- = \min(x_{ij}), j = 1, 2, \ldots, m \quad (6)$$

Distance to the ideal solution(D_j^+):

$$D_j^+ = \sqrt{\sum_{j=1}^{m} \left(x_{ij} - A_j^+\right)^2}, i = 1, 2, \ldots, n \quad (7)$$

Distance to the negative-ideal solution (D_j^-):

$$D_j^- = \sqrt{\sum_{j=1}^{m} \left(x_{ij} - A_j^-\right)^2}, i = 1, 2, \ldots, n \quad (8)$$

Calculating comprehensive attribute scores. The comprehensive molecular attribute score for each molecule

$$Score_i = \frac{D_j^-}{D_j^+ + D_j^-}, i = 1, 2 \ldots, n \quad (9)$$

Where:

$Score_i$ represents the comprehensive attribute score of molecule *i*.

In summary, CMAS could evaluate the overall performance of nucleoside hydrogels in terms of biocompatibility and anti-*P. gingivalis* activity.

Experimental verification

Characterization of nucleoside hydrogels

Hydrogel preparation procedure: 5.6 mg of the nucleoside derivative was dissolved in 0.2 mL ddH_2O and heated until a transparent liquid formed. Then, 0.2 mL of $AgNO_3$ solution (0.2 mol/L) was added and mixed thoroughly. The solution was cooled to room temperature and subjected to a tube-inversion test. Hydrogel formation was confirmed if no sample flow was observed upon inversion of the tube.

Scanning electron microscopy analysis: The hydrogel samples were freeze-dried to form xerogels, which were then attached to silica wafers and coated with gold. SEM images were obtained using an Inspect F50 scanning electron microscope (FEI, USA).

Fourier transform infrared spectroscopy: The hydrogel was freeze-dried to a xerogel, which was then analyzed using a Nicolet 6700 FTIR spectrometer (Thermo Fisher Scientific, USA) in the range of 400−4 000 cm^{-1}.

Powder X-ray diffraction: The hydrogel was freeze-dried to a xerogel, which was then analyzed using an X'Pert Pro MPD PXRD diffractometer (Netherlands). The PXRD patterns of hydrogel samples were recorded from 3° to 60° (2θ), with a current of 40 mA and a voltage of 40 kV.[28]

Rheological measurements of hydrogel: The hydrogel was heated to form a liquid and immediately transferred to the plate of an MCR302 rheometer (Anton Paar, Graz, Austria) for measurement.[29] The rheometer was equipped with a plate-plate geometry (25 mm diameter). The plate was preheated to 75 °C to prevent the samples from gelling before testing. After loading the samples, silicone oil was applied along the edge of the plate to prevent water evaporation during testing. A frequency sweep test was performed to obtain the elastic modulus (G') and loss modulus (G'') over the range of 0.1 to 100 rad/s at 25 °C, with a strain of 0.1%. A strain-dependent test was conducted with strain (γ) ranging from 0.1% to 100% at an angular frequency (ω) of 10 rad/s. The viscosity of the hydrogel was determined under shear strain with γ ranging from 0.1% to 100%. The self-healing ability of the hydrogel was assessed by alternating between γ values of 0.1% and 100% for four cycles.

Antibacterial activity assay for P. gingivalis. *P. gingivalis* (W83) was cultured at 37 °C in an atmosphere of 80% N_2, 10% H_2, and 10% CO_2. The bacterial culture medium consisted of brain-heart infusion (BHI) (BD Difco, USA), 1 μg/mL menadione, and 5 mg/L hemin (Sigma-Aldrich). A 1 mL bacterial suspension of *P. gingivalis*

(1×10^8 CFUs/ mL) was added to a 15 mL centrifuge tube, followed by the addition of nucleoside monomers and hydrogels at different concentrations. The bacterial suspensions were incubated for 24 h. Subsequently, the co-culture solution was diluted, and 100 μL of the diluted suspension was inoculated onto a BHI agar plate and incubated for approximately 3 days. Finally, the colony-forming units (CFUs) of *P. gingivalis* were counted.

Microscopy analysis for antibacterial property

Scanning electron microscopy analysis: A drop of *P. gingivalis* suspension was added to a clean glass slide and fixed with 2.5% glutaraldehyde for 4 h. After washing three times with PBS, the slides were gradually dehydrated using ethanol solutions with increasing concentrations (30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%) for 10 min each. The slides were then air-dried at room temperature and coated with gold. Finally, the specimens were observed by SEM (INSPECTF, FEI, USA) to visualize the surface morphology of *P. gingivalis*.

Confocal laser scanning microscopy (CLSM) analysis: *P. gingivalis* was stained using the Live/Dead BacLight Bacterial Viability Kit (L7012, Invitrogen, USA). Bacterial suspensions were collected and stained with SYTO 9 and propidium iodide (PI), then incubated in the dark for 15 min followed by a gentle rinse with PBS. A 5 μL aliquot of the stained bacterial suspension was dropped onto glass slides, and images were obtained using CLSM (FV3000, Olympus, Japan). The excitation/emission wavelengths were 488/530 nm for SYTO 9 and 561/630 nm for PI.

Cell counting kit-8 assays for biological toxicity. Cell counting kit-8 (CCK-8) assays were used to assess the toxicity of GMP, dGMP, and hydrogels on mouse skeletal muscle L929 cells. L929 cells were cultured in DMEM with 10% FBS and 1% penicillin-streptomycin at 37 °C in a 5% CO_2 atmosphere. Cells were seeded in 96-well plates at a density of 5×10^3 cells per well and cultured for 24 h. Nucleoside derivative solutions were then added to the cell culture medium and incubated for another 24 h. The biological toxicity of the samples was assessed using CCK-8, and the optical density (OD) values were measured at 450 nm using a microplate reader (SpectraMax iD5, China).[40]

Mouse model and experimental procedure. The mouse model used in this study was approved by the West China Hospital of Stomatology, Sichuan University (WCHSIRB-D-2024-420). All animal handling procedures adhered to the guidelines established by the Ethics Committee for animal care and use.

The periodontitis model was established using 10-week-old wild-type (WT) specific pathogen-free (SPF) female C57BL/6 mice (Taconic Farms, Model No. B6).[41] All mice had similar body weight, and were randomly assigned to each group. The mice were allowed to acclimate for 1 week before ligature placement, with regular access to rodent chow and water under conditions of 21 °C temperature, 50% humidity, and a 12-hour light/dark cycle. Anesthesia was induced using intraperitoneal injection of 10% chloral hydrate (4 mL/kg). Each mouse received treatment at only one periodontal site, and no animal was included in more than one experimental group. To ensure consistency across animals, the same ligature location, a 5-0 silk suture coated with *P. gingivalis* was ligated between the first and second maxillary molars to induce periodontitis.[30,31] Two models were established: an antibacterial treatment model, in which hydrogel was injected after successful model establishment, and an antibacterial prevention model, where hydrogel was injected concurrently with the start of the model. Euthanasia was performed using an overdose of anesthetic at the end of the experiments.[24,42]

Toxicity/biocompatibility evaluation: To assess toxicity, 100 μL of the target hydrogel (GMP 100 mmol/L, dGMP 100 mmol/L) was injected into the maxillary periodontal pockets, with PBS as a control. After 14 days, the mice were sacrificed, and tissues, including the heart, liver, spleen, lungs, and kidneys, were collected for H&E staining to evaluate in vivo toxicity. To evaluate biocompatibility, the target hydrogels were injected subcutaneously into the abdominal area of mice, with PBS as a control. After 7 days, subcutaneous tissues were collected for H&E staining to assess biocompatibility.

Antibacterial treatment model for periodontitis: This model included eight groups ($n = 3$ per group): a periodontitis group, three control groups (silver ion group, GMP monomer group, dGMP monomer group), and four hydrogel groups with varying concentrations (GMP 50 mmol/L, GMP 100 mmol/L, dGMP 50 mmol/L, dGMP 100 mmol/L). In total, 24 mice were randomly assigned to the groups. Fourteen days post-induction of periodontitis, 100 μL of the respective material was injected into the periodontal pockets. Ligatures were examined every 3 days, and replaced if loosened or detached. After another 14 days, periodontal tissue from the maxillary molar area was collected for H&E, and IHC (TNF-α and TGF-β).[24] Maxillary bone tissue was collected for micro-CT analysis.[24] Additional analysis information could be found in Supplementary Method 3.

Antibacterial prevention model for periodontitis: This model included five groups: a periodontitis group, a healthy control group, and two hydrogel groups (GMP 100 mmol/L, dGMP 100 mmol/L). A total of 15 mice were randomly assigned to these groups ($n = 3$ per group). Following induction of periodontitis, 100 μL of the corresponding hydrogel was injected into the periodontal pockets the following day. After 14 days, periodontal tissue from the maxillary molar area was collected for H&E, and IHC staining (TNF-α and TGF-β), while maxillary bone tissue was collected for micro-CT analysis.

Data analysis. Quantitative results are expressed as mean ± standard error (SE). Statistical analysis was conducted using GraphPad Prism, with each experiment repeated at least three times. One-way analysis of variance (ANOVA, Dunnett's multiple comparisons test) was employed for data evaluation. Statistical significance was defined as $P < 0.05$: *, $P < 0.01$: **, $P < 0.001$: *** and $P < 0.000\ 1$: ****.

Ethical approval: The mouse model used in this study was approved by the West China Hospital of Stomatology, Sichuan University (WCHSIRB-D-2024-420). All animal handling procedures adhered to the guidelines established by the Ethics Committee for animal care and use.

DATA AVAILABILITY

All data supporting the key findings of this study are provided. Existing datasets analyzed and those generated during the study are available on GitHub (https://github.com/leescu/NBPND). Source data have been deposited in Zenodo: https://doi.org/10.5281/zenodo.10723552. All code used in this study is available in the GitHub repository at https://github.com/leescu/NBPND (https://doi.org/10.5281/zenodo.10723747). The MBSI and CMAS allowing analysis with either the data provided in this study or other independent datasets are available on GitHub (https://github.com/leescu/NBPND/MBPA).

MATERIALS AVAILABILITY

Supplementary Data 1–10 supporting this study are available at zenodo (NBPND, https://doi.org/10.5281/zenodo.18487279). All data supporting the key findings of this study are provided. Existing datasets analyzed and those generated during the study, are available on GitHub (https://github.com/leescu/NBPND). Source data have been deposited in Zenodo (https://doi.org/10.5281/zenodo.10723552). All code used in this study is available in the GitHub repository at https://github.com/leescu/NBPND (https://doi.org/10.5281/zenodo.10723747). The MBSI and CMAS allow analysis with either the data provided are available on GitHub (https://github.com/leescu/NBPND).

ACKNOWLEDGEMENTS

National Key R&D Program of China (No. 2022YFC2402901 to H.Z.); National Natural Science Foundations of China (No. 82571160 to H.Z. and No. 82370962 to H.X.); Central Government Guided Local Science and Technology Development Fund Projects (No. 2024ZYD0176 to H.Z.); Research Funding from State Key Laboratory of Oral Diseases, West China School/Hospital of Stomatology Sichuan University (No. RCDWJS2026-14, SKLOD-2025KP001, SKLOD-R011 to H.Z.); The Natural Science Foundation of Sichuan Province of China (No. 2026NSFSC1587 to W.L. and 2026NSFSC0489 to H.X.); China Postdoctoral Science Foundation (CPSF, No. 2025M781704 to W.L.), the Postdoctoral Fellowship Program of CPSF (No. GZB20250442 to W.L.). The authors thank Kaichao Wang and Lideng Cao (State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University) for helping us perform studies. The authors also thank Siyuan Hao for helping with data collation.

AUTHOR CONTRIBUTIONS

Conceptualization: Q.C., H.X., H.Z. Methodology: W.L., Y.W. Investigation: W.L., Y.W., Z.H. Visualization: W.L. Supervision: H.X., H.Z. Writing—original draft: W.L., Y.W., Z.H. Writing—review & editing: F.Z., F.S., Y.Y.W.L., and Y.W., and Z.H. contributed equally to this work. Correspondence to H.X. and H.Z.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41368-026-00438-3.

Competing interests: The authors declare no competing interests.

REFERENCES

1. Peters, G. M. et al. G4-quartet·M(+) borate hydrogels. *J. Am. Chem. Soc.* **137**, 5819–5827 (2015).
2. Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci.* **48**, 2013–2018 (1962).
3. Li, Y. et al. A guanosine-quadruplex hydrogel as cascade reaction container consuming endogenous glucose for infected wound treatment—a study in diabetic mice. *Adv. Sci.* **9**, 2103485 (2022).
4. Wang, Z. et al. High-strength and injectable supramolecular hydrogel self-assembled by monomeric nucleoside for tooth-extraction wound healing. *Adv. Mater.* **34**, 2108300 (2022).
5. Liu, T. et al. pH-responsive dual-functional hydrogel integrating localized delivery and anti-cancer activities for highly effective therapy in PDX of OSCC. *Mater. Today* **62**, 71–97 (2023).
6. Qi, J. et al. Inosine-based supramolecular hydrogel for highly efficient PD-L1 blockade therapy via mediating CD8+ T Cells. *Adv. Funct. Mater.* **32**, 2204273 (2022).
7. Ramin, M. A. et al. Cation tuning of supramolecular gel properties: a new paradigm for sustained drug delivery. *Adv. Mater.* **29**, 1605227 (2017).
8. Peters, G. M. & Davis, J. T. Supramolecular gels made from nucleobase, nucleoside and nucleotide analogs. *Chem. Soc. Rev.* **45**, 3188–3206 (2016).
9. Adams, D. J. Personal perspective on understanding low molecular weight gels. *J. Am. Chem. Soc.* **144**, 11047–11053 (2022).
10. Li, W. et al. Developing a machine learning model for accurate nucleoside hydrogels prediction based on descriptors. *Nat. Commun.* **15**, 2603 (2024).
11. Catacutan, D. B., Alexander, J., Arnold, A. & Stokes, J. M. Machine learning in preclinical drug discovery. *Nat. Chem. Biol.* **20**, 960–973 (2024).
12. Schenone, M., Dančík, V., Wagner, B. K. & Clemons, P. A. Target identification and mechanism of action in chemical biology and drug discovery. *Nat. Chem. Biol.* **9**, 232–240 (2013).
13. Peteani, G., Huynh, M. T. D., Gerebtzoff, G. & Rodríguez-Pérez, R. Application of machine learning models for property prediction to targeted protein degraders. *Nat. Commun.* **15**, 5764 (2024).
14. Stokes, J. M. et al. A Deep learning approach to antibiotic discovery. *Cell* **180**, https://doi.org/10.1016/j.cell.2020.01.021 (2020).
15. Wang, Y. et al. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* **37**, W623–W633 (2009).
16. Altae-Tran, H., Ramsundar, B., Pappu, A. S. & Pande, V. Low data drug discovery with one-shot learning. *ACS Cent. Sci.* **3**, 283–293 (2017).
17. Huang, R. et al. Modelling the Tox21 10 K chemical profiles for in vivo toxicity prediction and mechanism characterization. *Nat. Commun.* **7**, 10425 (2016).
18. Enot, D. P. et al. Preprocessing, classification modeling and feature selection using flow injection electrospray mass spectrometry metabolite fingerprint data. *Nat. Protoc.* **3**, 446–470 (2008).
19. Artrith, N. et al. Best practices in machine learning for chemistry. *Nat. Chem.* **13**, 505–508 (2021).
20. Li, W. et al. Periodontitis and the risk of all-cause and cause-specific mortality among US adults with diabetes: a population-based cohort study. *J. Clin. Periodontol.* **51**, 288–298 (2024).
21. Kinane, D. F., Stathopoulou, P. G. & Papapanou, P. N. Periodontal diseases. *Nat. Rev. Dis. Prim.* **3**, 17038 (2017).
22. Luo, L.-S. et al. The spatial and temporal trends of severe periodontitis burden in Asia, 1990-2019: a population-based epidemiological study. *J. Periodontol.* **93**, 1615–1625 (2022).
23. Pihlstrom, B. L., Michalowicz, B. S. & Johnson, N. W. Periodontal diseases. *Lancet* **366**, 1809–1820 (2005).
24. Xu, Y. et al. Microenvironment-responsive metal-phenolic nanozyme release platform with antibacterial, ROS scavenging, and osteogenesis for periodontitis. *ACS Nano* **17**, 18732–18746 (2023).
25. Xu, S. et al. Alleviate periodontitis and its comorbidity hypertension using a nanoparticle-embedded functional hydrogel system. *Adv. Health. Mater.* **12**, e2203337 (2023).
26. Hafezalkotob, A., Hafezalkotob, A., Liao, H. & Herrera, F. Interval MULTIMOORA method integrating interval Borda rule and Interval Best-Worst-Method-based weighting model: case study on hybrid vehicle engine selection. *IEEE Trans. Cyber.* **50**, 1157–1169 (2020).
27. Pillinger, T. et al. Antidepressant and antipsychotic side-effects and personalised prescribing: a systematic review and digital tool development. *Lancet Psychiatry* **10**, 860–876 (2023).
28. Matsumura, K., Tateno, K., Tsuchido, Y. & Kawai, H. Spacer-dependent cooperativity of helicity in fluorescent bishelical foldamers based on L-shaped dibenzopyrrolo[1,2-a][1,8]naphthyridine. *ChemPlusChem* **86**, 1421–1425 (2021).
29. Zhao, H. et al. Dual-functional guanosine-based hydrogel integrating localized delivery and anticancer activities for cancer therapy. *Biomaterials* **230**, 119598 (2020).
30. Roldan, L. et al. A novel injectable piezoelectric hydrogel for periodontal disease treatment. *ACS Appl. Mater. Interfaces.* **15**, 43441–43454 (2023).
31. Li, J. et al. 2D MOF periodontitis photodynamic ion therapy. *J. Am. Chem. Soc.* **143**, 15427–15439 (2021).
32. Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. in *Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2623–2631 (ACM, 2019).
33. Figueiredo, M. A. T., Nowak, R. D. & Wright, S. J. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE J. Sel. Top. Signal Process.* **1**, 586–597 (2007).
34. Quinlan, J. R. Induction of decision trees. *Mach. Learn.* **1**, 81–106 (1986).
35. Yu, H.-F., Huang, F.-L. & Lin, C.-J. Dual coordinate descent methods for logistic regression and maximum entropy models. *Mach. Learn.* **85**, 41–75 (2011).
36. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).
37. Chen, T. & Guestrin, C. in *Proc.22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (Association for Computing Machinery, 2016).
38. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
39. Lee, P. in *Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* 3–4 (Association for Computing Machinery, Anchorage, 2019).
40. Li, T. et al. Super-rapid in situ formation of a silver ion-induced supramolecular hydrogel with efficient antibacterial activity for root canal disinfection. *ACS Appl. Mater. Interfaces.* **15**, 29854–29865 (2023).
41. Marchesan, J. et al. An experimental murine model to study periodontitis. *Nat. Protoc.* **13**, 2247–2267 (2018).
42. Guo, C. et al. Engineering magnetically induced antibacterial organic/inorganic hybrid nanoparticles for the treatment of periodontitis. *Giant.* **13**, 100145 (2023).