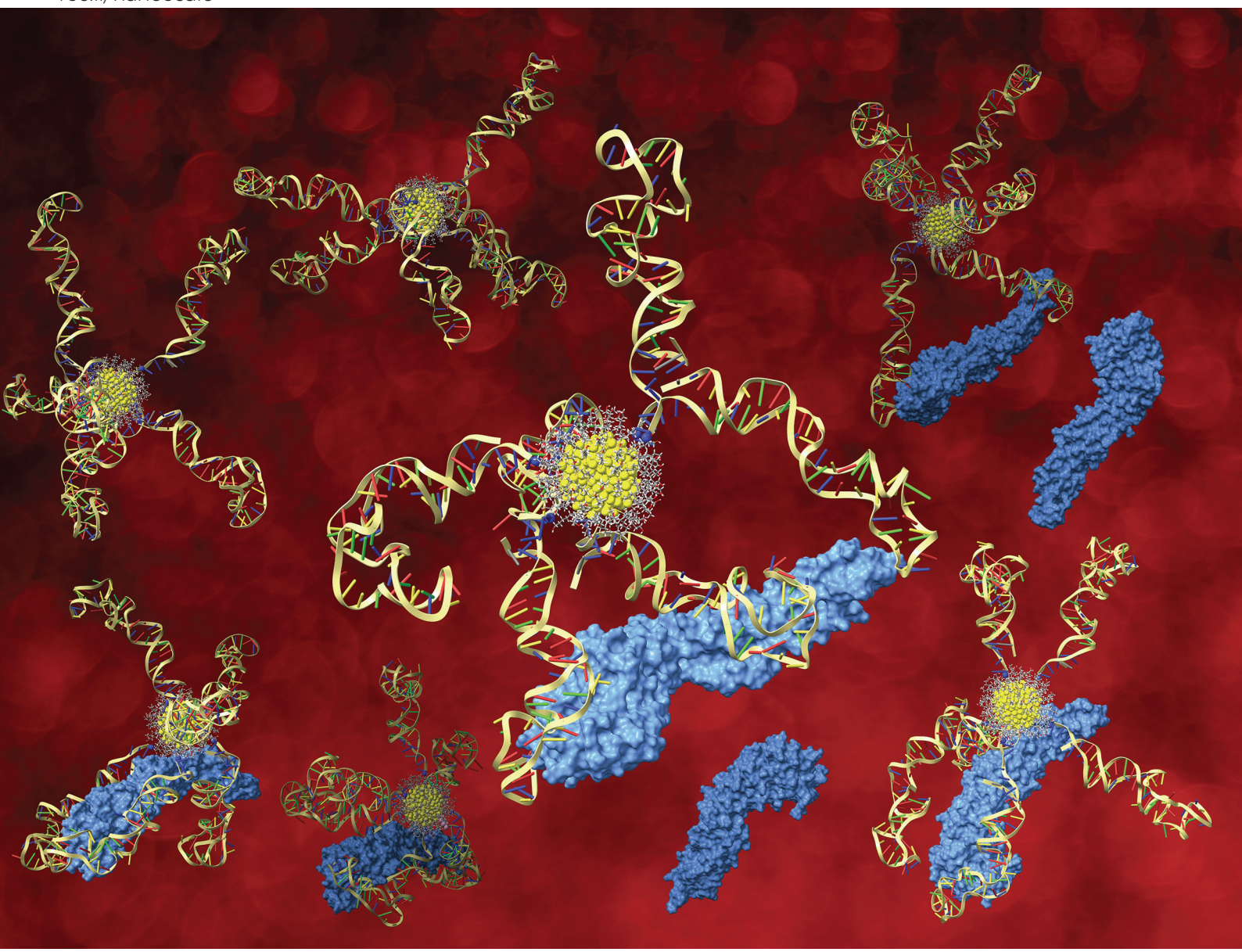
Volume 18
Number 2
15 January 2026
Pages 523-1106

Nanoscale

rsc.li/nanoscale



ISSN 2040-3372

ROYAL SOCIETY OF CHEMISTRY

PAPER
Alessandro Mossa and Giorgia Brancolini
Rational design of gold nanoparticles functionalized with aptamers for improved West Nile virus detection

NCNST

Nanoscale

PAPER

View Article Online
View Journal | View Issue

Cite this: *Nanoscale*, 2026, **18**, 723

Rational design of gold nanoparticles functionalized with aptamers for improved West Nile virus detection

Alessandro Mossa * and Giorgia Brancolini *

Received 30th July 2025,
Accepted 21st November 2025

DOI: 10.1039/d5nr03228h

rsc.li/nanoscale

This study employs a multiscale approach to investigate the interaction dynamics between the West Nile virus (WNV) envelope (E) protein and gold nanoclusters functionalized with DNA aptamers. By integrating aptamer secondary and tertiary structure prediction, structural docking and atomistic molecular dynamics simulations, we reveal the structural, dynamic and electrostatic factors governing the aptamer-target WNV E protein recognition at the nanoscale. Two gold nanoclusters, $Au_{144}(SR)_{60}$ and $Au_{314}(SR)_{96}$ (SR = thiolate ligand), were examined to assess how nanoparticle size and surface functionalization influence aptamer anchoring and binding stability under physiological salt conditions. This approach enhances aptamer functionalization strategies for detecting West Nile virus and creates a flexible framework for aptamer-based diagnostics for other emerging pathogens, with implications for diagnostics across a range of viral and protein biomarkers.

Introduction

The clinical application of nanotechnology in biosensing presents an exciting frontier for improving the detection of viral biomarkers in complex biological fluids, such as blood and saliva. Achieving high sensitivity and specificity is crucial for effective diagnostics, particularly for pathogens that pose significant public health risks. Among these pathogens, the West Nile virus (WNV) is a noteworthy concern due to its association with severe neurological diseases, including encephalitis and flaccid paralysis. The limitations of existing diagnostic methods for WNV, which often require sophisticated facilities and trained personnel, underscore the urgent need for innovative and accessible detection strategies.[1,2]

The interaction between proteins and functionalized gold nanoparticles (AuNPs) plays a fundamental role in optimizing surface properties necessary for biosensor applications.[3] Utilizing monolayer-coated AuNPs that incorporate DNA aptamers as functional groups allows fine-tuning of surface characteristics to enhance interactions with biological targets.[4] In particular, controlling the density, orientation, and binding affinity of biomolecules, such as proteins, is critical to improve the performance of monolayer-coated AuNP systems.[5]

This study aims to address a fundamental gap by developing label-free plasmonic nanoparticles (NPs)-based bioactive platforms for the spectroscopic recognition and quantification of pathogens. Specifically, it focuses on the WNV's envelope (E) protein, a key glycoprotein that plays a pivotal role in the virus's pathogenesis,[6] thus advancing our understanding of virus-nanomaterial interactions at the molecular level.

Although substantial progress over the past decade has advanced our understanding of sensing methods to detect viruses belonging to the flaviviridae family,[7] such as Dengue virus (DENV), Zika virus (ZIKV), Yellow Fever virus (YFV), and Tick-borne encephalitis virus (TBEV), methods for detecting West Nile virus (WNV) using nanoparticle-based nanosensors remain underdeveloped. Currently, the available diagnostic techniques for DENV and WNV primarily rely on cell culture,[8] polymerase chain reaction (PCR),[9] and Enzyme-Linked Immunosorbent Assay (ELISA).[10,11] These conventional diagnostic methods often necessitate sophisticated medical facilities and trained personnel, which are not suitable for resource-limited settings.

The development of nanoparticle-based biosensors represents a promising approach by exploiting the fundamental physicochemical interactions between nanoparticles and biomolecules to enhance sensitivity and specificity in viral detection. Recent advances in aptamer-based biosensors, or aptasensors, highlight the importance of understanding these interactions at the molecular level, providing a powerful platform for disease diagnostics through precise and selective biomolecular recognition.[12] Aptamers, which are single-stranded oligonucleotides, possess several advantages including high stability, ease of chemical functionalization, low cost, and minimal batch-to-batch variation.[13]

Notably, colorimetric-based aptasensors have emerged as favorable analytical methods, utilizing nanomaterials such as

Center S3, CNR Institute of Nanoscience, via Campi 213/A, 41125 Modena, Italy.
E-mail: alessandro.mossa@nano.cnr.it, giorgia.brancolini@nano.cnr.it

nanozymes and plasmonic nanoparticles for enhanced signal detection.[14,15] However, electrochemical-based aptasensors have gained dominance due to their superior signal transduction capabilities.[12,16] Despite these advancements, existing methods often suffer from a lack of signal amplification, leading to sensitivity limitations.[17]

In recent studies, novel aptamer designs have shown promise for specific viral detection. For example, Lee and Zeng[14] developed Zika NS1-binding aptamers with high affinity, demonstrating their potential for sensitive viral detection in ELISA-based assays. Similarly, Matsunaga *et al.*[15] utilized DNA aptamers containing unnatural bases to enhance specificity for detecting Dengue virus NS1 serotypes. Morais *et al.*[17] successfully selected single-stranded DNA (ssDNA) aptamers specific to the Zika virus through whole-virus Systematic Evolution of Ligands by Exponential Enrichment (SELEX),[18] allowing Zika to be differentiated from Dengue virus. In this work, the term single-stranded oligonucleotides refers to short, single-stranded nucleic acid sequences. ssDNA denotes oligonucleotides composed exclusively of deoxyribonucleotides, typically used as the structural basis for DNA aptamers. The term aptamer is reserved for a subset of ssDNA (or RNA) sequences that, through the SELEX process, fold into defined three-dimensional conformations capable of specifically recognizing and binding to a target molecule with high affinity.

Aptamers act as synthetic analogs of antibodies, exhibiting high affinity and specificity toward a wide variety of targets, including proteins, viruses, and small molecules. Emerging methodologies incorporating artificial intelligence (AI) have further streamlined the aptamer design process, enhancing the efficiency of SELEX.[19,20] By integrating computational techniques such as molecular dynamics simulations and machine learning, researchers can optimize aptamer-target binding predictions and improve the overall design of aptamers for diagnostic applications.

In recent years, computational modeling—particularly molecular dynamics (MD) simulations—has become a powerful complement to experimental approaches for elucidating aptamer behavior in biosensing systems. A typical *in silico* workflow involves aptamer structure prediction, target docking, and MD refinement to evaluate the stability and binding energetics of aptamer–ligand complexes.[21]

Several MD studies have provided detailed insight into aptamer–nanomaterial and aptamer–protein interactions relevant to biosensing.[22,23] For instance, simulations of RNA aptamers immobilized on thiol-functionalized gold surfaces have revealed how spacer length, aptamer orientation, and interfacial ion distributions affect target recognition efficiency.[4] Similarly, atomistic simulations of aptamer–protein complexes, such as anti-MUC1 systems, have shown how subtle sequence or structural variations modulate binding affinity and conformational stability.[24] Other examples include MD analyses of ssDNA aptamers adsorbed on citrate-capped gold nanoparticles, highlighting the electrostatic nature of aptamer adsorption and the influence of divalent ions on complex formation.[25] These studies collectively demonstrate the central role of MD simulations in connecting nanoscale structure, surface chemistry, and binding performance, thereby guiding the rational design of high-efficiency aptasensors.

In light of recent advances by Park *et al.*,[26] who experimentally demonstrated the Tr-WNV-37 truncated aptamer in a rapid, cost-effective electrochemical biosensor for West Nile virus (WNV) detection, we use the Tr-WNV-37 sequence as the foundation for a computational exploration of aptamer performance when anchored to gold nanoclusters. We design and optimize high-affinity DNA aptamers targeting the WNV envelope (E) protein by predicting their secondary and tertiary structures, linking them *via* a T_6 spacer, and modeling their immobilization on hexanol-thiol-functionalized gold nanoclusters. This computational approach enables the identification of key structural determinants governing aptamer binding and provides insights that complement and extend the experimental findings.

Using state-of-the-art tools we perform *in silico* simulations to enhance binding specificity. Selected aptamer sequences are anchored onto hexanol-thiol-functionalized gold nanoclusters of varying sizes to assess their stability and binding affinity.

Our methodology involves the development of tailored force field parameters to simulate these systems accurately,[27] revealing that an increase in aptamer density on $Au_{314}(SR)_{96}$ enhances additional interactions without disrupting primary binding, whereas $Au_{144}(SR)_{60}$ shows less promising stability. Thiol angle analysis highlights variability in anchoring orientations due to nanoparticle diffusion, suggesting a flexible and robust binding interface between the aptamer and E protein.

The computational findings are further supported by free energy calculations, demonstrating a lower ΔG of binding upon the aptamer anchoring to gold nanoclusters, which is indicative of a more stable aptamer-protein interaction. This approach not only validates the experimental stability observed in Park *et al.*'s work,[26] but also highlights the fact that the aptamer immobilization on surfaces does not weaken its binding affinity for the target protein.

To summarize, the aim of the present work is to develop a computational strategy to engineer gold nanoparticle (AuNP), aptamer conjugates that maintain and potentially enhance binding affinity under high ionic strength conditions, environments that typically compromise biosensor performance.[3,27] Through advanced molecular dynamics simulations, we present the first detailed characterization of the structural dynamics, surface anchoring behavior, and electrostatic interactions within the Tr-WNV-37–AuNP–protein complex. This theoretical framework elucidates the molecular determinants of aptamer functionality in realistic biosensing settings and establishes design principles for next-generation aptamer-based nanobiosensors targeting WNV and related flaviviruses.

2 Methodology

2.1 Initial structures

The structure of the envelope glycoprotein (E) of the West Nile virus has been resolved by X-ray diffraction and has been de-

posited as PDB ID: 2HG0.[28] For our simulations, we replace the expression tag "HHHHHH" at positions 403–408 with a single serine (the complete sequence is reported in section S1 in SI). The E protein has a three-domain structure represented in Fig. 1(A). Domain I (DI) is an eight-stranded central β-barrel that is connected by hinge regions to domain II (DII), which contains at the distal end the putative fusion loop, and domain III (DIII), an immunoglobulin-like fold thought to contain the receptor binding site.[29] E-protein dimers cover the whole lipid membrane of the virion in a smooth, quasi-icosahedral shell.

The protonation state of the E protein was computed using the web server *H*++[31] under the conditions 150 mM l^{-1} NaCl, pH 7.4, external and internal dielectric constants valued 80 and 10, respectively. The total charge of the E protein is then $-2e$, its mass 42.1 kDa. All simulations were conducted at a physiological ionic strength of 150 mM, consistent with the salt concentrations typically used in aptamer folding buffers,[26] and biosensing assays.[3,27]

The structures of gold nanoparticles $Au_{144}(SR)_{60}$[32] and $Au_{314}(SR)_{96}$[33] have been retrieved from the webserver NanoModeler 2.0,[34] which was also used to coat the NPs with thiolate ligands $SC_6H_{12}OH$. The use of such self-assembled monolayers (SAMs) in electrochemical biosensors has been shown to improve both the formation and stability of aptamer–gold interfaces, thus improving sensor performance.[35] The functionalization with 6-mercapto-1-hexanol self-assembled monolayer (SAM) was chosen for its chemical stability and its widespread use as the preferred functionalization in gold electrochemical biosensors. This selection is supported by multiple studies,[22,36] demonstrating the effectiveness of functionalized gold surfaces in biosensing applications *via* atomistic simulations. In particular, in our recent work we discussed the role of SAMs in stabilizing nanostructured interfaces,[35] by performing atomistic simulations compared with SPR experiments. Moreover, Imanparast *et al.*[37] showed that 6-mercapto-1-hexanol facilitates functionalization while preserving structural stability and enabling targeted conjugation strategies. Furthermore, Ruan *et al.*[4] conducted molecular dynamics simulations of an RNA aptasensor immobilized on a hexanol-thiol-functionalized gold surface, demonstrating that the aptamer remains oriented parallel to the surface, preserving its structural integrity and binding affinity–features critical for biosensing performance.

2.1.1 Structure of the Tr-WNV-37 aptamer. The initial phase of the study involved the design and structural modeling of aptamers targeting West Nile virus (WNV) proteins. For the WNV envelope (E) protein, the aptamer candidate Tr-WNV-37 was selected based on prior literature[26] where it was identified through SELEX and evaluated for its predicted binding affinity and structural compatibility.

Special attention was paid to domain III (DIII) of the E protein, a well-characterized epitope known to be targeted by neutralizing antibodies.[38] Aptamer design was directed specifi-

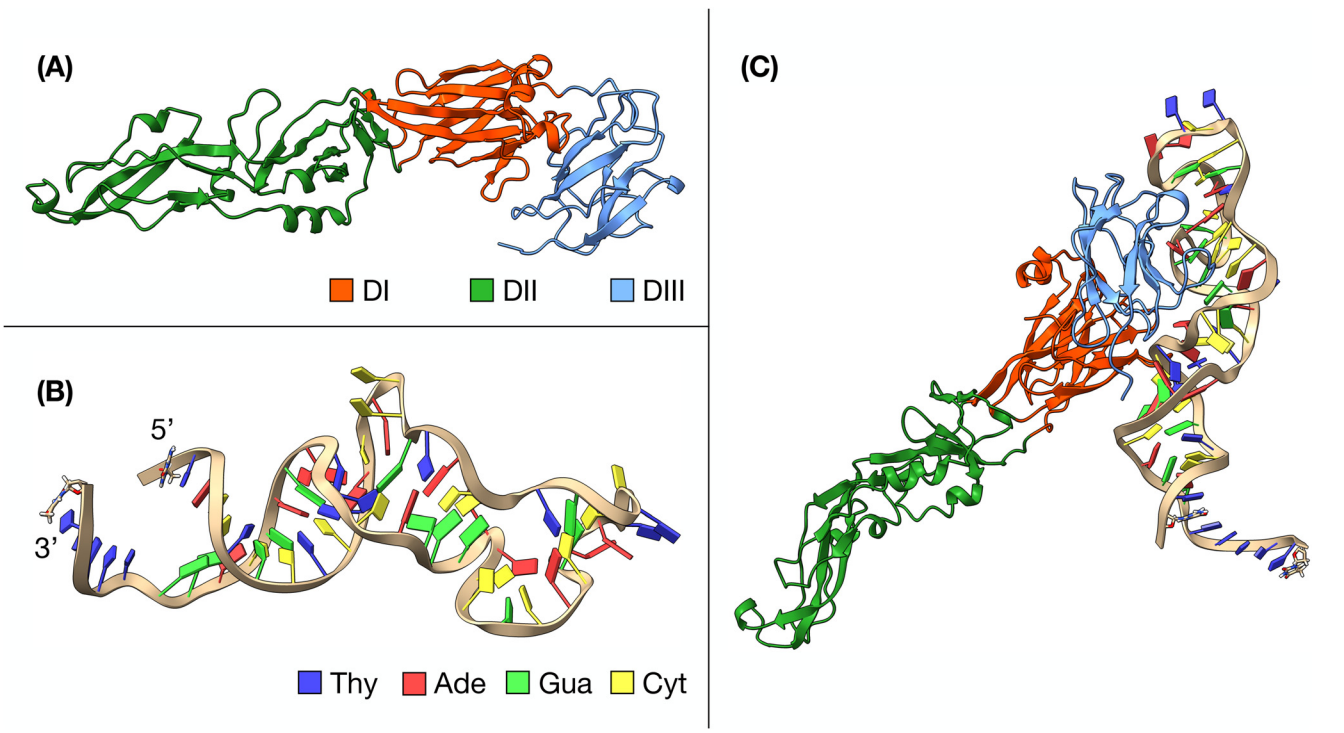


Fig. 1 (A) Three-domain structure of the E protein of the West Nile virus. (B) Atomic model of the aptamer Tr-WNV-37 used in the docking procedure. (C) Best-scoring aptamer-protein complex as determined by docking. All images of molecular models in this paper have been created with UCSF ChimeraX.[30]

cally toward this region to ensure high-affinity and selective binding.

The modeling pipeline begins with UNAfold,[39] which predicts the most thermodynamically stable 2D secondary structure of the aptamer based on its sequence. This structure is then inputted into RNAComposer[40,41] to generate a 3D atomic model. The resulting RNA model is converted into its DNA counterpart (Fig. 1(B)), energy-minimized, and subsequently used for docking simulations against the E protein. Docking allows for the identification of the most favorable binding orientation and key interaction sites between the aptamer and its target.

One may wonder if there are quicker alternatives to the time-consuming modeling pipeline describe above. Even though some measure of success has recently been reported[42] on using directly AlphaFold 3[43] to predict aptamer 3D structure, this particular sequence belongs to the class of synthetic oligonucleotides underrepresented in the databases used for training, with the result that the predicted structure has in this case very low confidence and is in fact hardly believable in its attempt to reduce the aptamer's shape into a regular double-helix.

Pons *et al.*[44] demonstrated, using surface plasmon resonance (SPR), that the introduction of a spacer group during aptamer immobilization significantly influences the sensor response. Specifically, they observed that a spacer length of six thymidines (T_6) enhanced aptamer–target affinity compared to no spacer (T_0), as evidenced by a marked increase in SPR signal amplitude. This improvement was attributed to the spacer's ability to reduce steric hindrance and promote an optimal orientation of the aptamer on the sensor surface, thereby facilitating target accessibility and binding efficiency. Guided by these findings, we incorporated a T_6 spacer in our design to achieve a similar effect–enhancing aptamer–target interactions by minimizing surface-induced conformational constraints. This choice is further supported by the molecular dynamics simulations of Ruan *et al.*,[4] which showed that a T_6 spacer preserves the aptamer's structural integrity and orientation upon immobilization, thereby maintaining its binding affinity. The rationale for employing the T_6 spacer is thus firmly grounded in established principles of aptamer immobilization and SPR-based biosensing.

The sequence of the aptamer Tr-WNV-37 modified by the insertion of six thymines at the 3′ end is 5′-GTACC TGCTC TACAC CGTAC CTGCT CTAAG CACGC CAAGG GACTA TAGAG CAGGT TTTTT-3′.

2.1.2 Docking of the aptamer on the E protein. Docking simulations were performed using HADDOCK 2.2,[45] following the default protein–DNA docking protocol as described by Van Dijk *et al.*[46]

The Tr-WNV-37 aptamer and the West Nile virus (WNV) envelope (E) protein were used as docking partners. Domain III (DIII), a well-characterized target known for its high antigenicity and critical role in receptor binding, has been extensively studied as a binding epitope for both aptamers and antibodies.[28] Residues on domain III (DIII) previously identified as directly involved in aptamer binding were defined as active, meaning that the scoring algorithm assigns a penalty if these residues are not interacting within the docked configuration. Residues from domains I and II (DI and DII) were defined as passive: they were included in the docking calculations and could contribute to the interaction through steric and non-bonded effects, but the scoring function does not enforce their participation. Importantly, DI and DII were not truncated or removed from the protein structure prior to docking. For the aptamer, all nucleotides were marked as active to enable full exploration of possible binding interfaces.

Rigid-body docking was carried out to efficiently explore initial binding orientations according to a well-established computational protocol,[36,47,48] followed by enhanced-sampling MD refinement to incorporate the intrinsic flexibility of both the aptamer and the protein at the interface. When rigid-body docking was carried out, it generated multiple aptamer–protein complexes. Despite differences in the number of top-ranked solutions per run, the 6 best models were selected based on their HADDOCK scores. These models were further evaluated by cluster analysis, ranking them according to their *Z*-scores and energy terms.

The most representative cluster was selected based on its relative population, accounting for 60% of all models (Cluster 1), indicating it was the most recurrent among the generated solutions. A second cluster was found to account for 20% of the resulting poses (Cluster 2), while the remaining association complexes were distributed among four different outlier clusters (Clusters 3, 4, 5, and 6), each representing only 5% of the structures. Additionally, the most representative cluster had the lowest *Z*-score (−2.1). This combination of high recurrence and favorable scoring suggests that this cluster represents a statistically robust and high-quality docking solution. The energy breakdown revealed that the primary driving force for the initial association complex was electrostatics, followed by van der Waals interactions. The representative structure corresponding to this most favorable *Z*-score and dominant population was chosen as the starting configuration for molecular dynamics (MD) simulations (see Fig. 1(C)), ensuring a stable and energetically favorable model for further refinement and analysis.

The hydrogen bonds present in the starting aptamer-protein configuration are detailed in Table 1. They can be compared with the map of most frequent contacts in Fig. 16 to assess that most (but not all) of the initial interactions are preserved during the simulation, a signal that the initial docking

Table 1 Hydrogen bonds in the best-scoring aptamer-protein complex as determined by docking

Type	Id	Type	Id	Notes
SER	16	Cyt	10	
LEU	349	Cyt	20	
LEU	349	Cyt	21	
THR	350	Cyt	20	
PHE	379	Cyt	13	
ARG	388	Thy	22	Double
ARG	388	Ade	38	
GLN	392	Ade	38	

analysis is sound, and that MD has had enough time to correct the picture, when needed.

2.1.3 Simulated systems. The selection of the systems to be simulated was based on two principles: (1) studying two gold NPs of different size to begin exploring possible size and curvature effects; (2) testing varied aptamer densities. The choice of 2 and 4 aptamers on the smaller NP, and 3 and 6 aptamers on the larger NP is meant to obtain a similar aptamer density. All systems prepared for simulation are listed in Table 2, and graphically represented in Fig. 2. The inclusion of gold NPs without aptamers is mainly motivated by the need to obtain a reference against which to evaluate the soundness of the diffusion coefficients estimated in section 3.3.

All systems in Table 2, except the two consisting of NPs without aptamers, have also been simulated in the presence of the E protein, always starting from a configuration in which the relative position between the protein and one of the aptamers is initially the same resulting from the docking procedure detailed in section 2.1.2. In all of the MD traces this first aptamer remains close to the E protein throughout the entire simulation, even if a rearrangement of their relative position is quite common. If presents, also other aptamers can, and often do, get close to the protein, either for a while, or until the end of the trajectory. A representative final frame from each of these simulations is depicted in Fig. 3. An additional set of simulations was run for the aptamer alone (not connected to any NP) in interaction with the E protein.

Table 2 Technical details about the systems simulated in the absence of the E protein, illustrated in Fig. 2

Data set	Mass [kDa]	Charge [e]	Apt. density [nm^{-2}]
$Au_{314}(SR)_{96}$	74.6	0	0.0000
$Au_{314}(SR)_{96}$ + 1 apt.	92.9	−59	0.0468
$Au_{314}(SR)_{96}$ + 3 apt.	125.5	−177	0.1404
$Au_{314}(SR)_{96}$ + 6 apt.	184.5	−354	0.2808
$Au_{144}(SR)_{60}$	36.4	0	0.0000
$Au_{144}(SR)_{60}$ + 1 apt.	54.7	−59	0.0719
$Au_{144}(SR)_{60}$ + 2 apt.	70.3	−118	0.1438
$Au_{144}(SR)_{60}$ + 4 apt.	104.2	−236	0.2876

2.2 Force field

In a series of preliminary tests, the force field that better preserved the structure of the aptamer turned out to be amber99bsc1,[49] an adaptation for GROMACS of the Amber99 force field Parmbsc1 for DNA.[50] We availed ourselves of the webserver NanoModeler 2.0[34] to join the gold nanoparticles $Au_{144}(SR)_{60}$[32] and $Au_{314}(SR)_{96}$[33] with the thiolate ligands, which we modeled using R.E.D. Server Development 2.0.[51]

Once the complete aptamer–SAM–AuNP system was constructed using an in-house script, force-field parameters were derived to ensure full compatibility with amber99bsc1. RESP charges for the linker atoms were derived at the HF/6-31G* level of theory using the PyRED server,[51] following the RESP-A1 scheme, which is fully consistent with AMBER standards. The derivation workflow, comprising geometry optimiz-

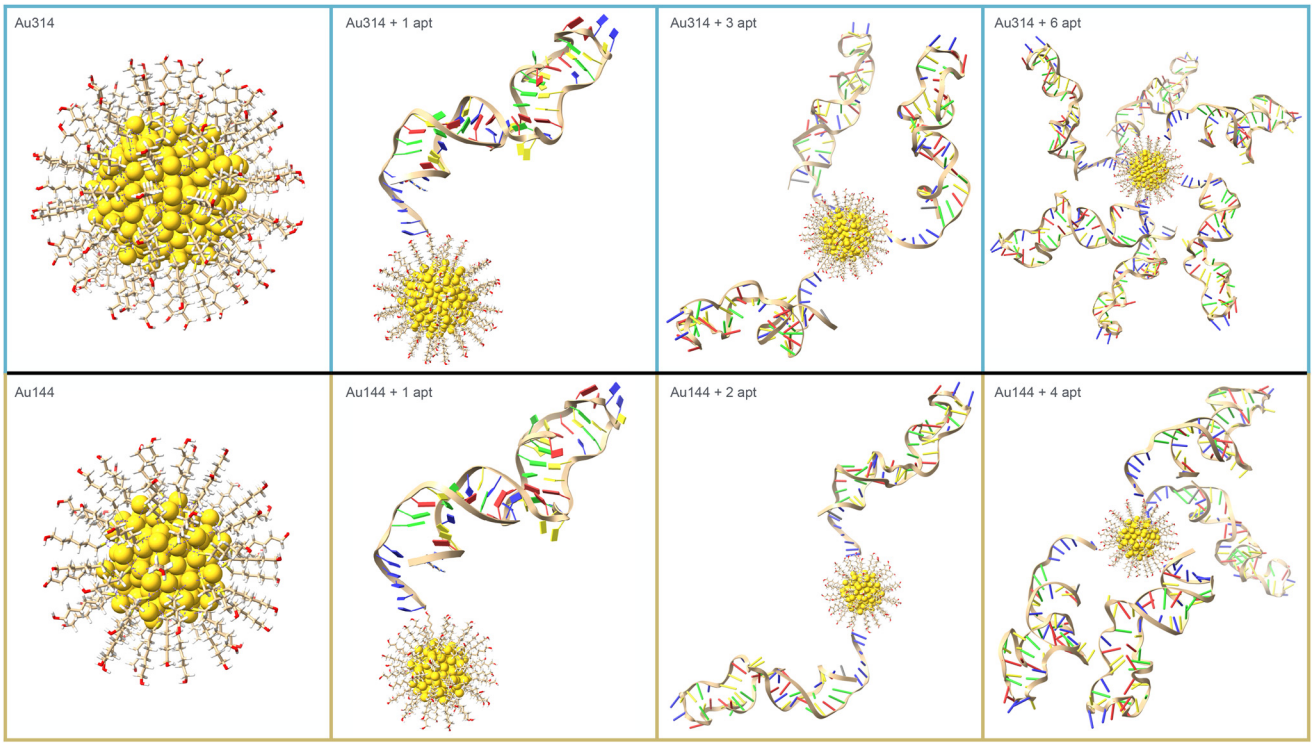


Fig. 2 Initial configuration of the systems that have been simulated in the absence of the E protein. The number of aptamers has been chosen to have matching aptamer density (see Table 2).

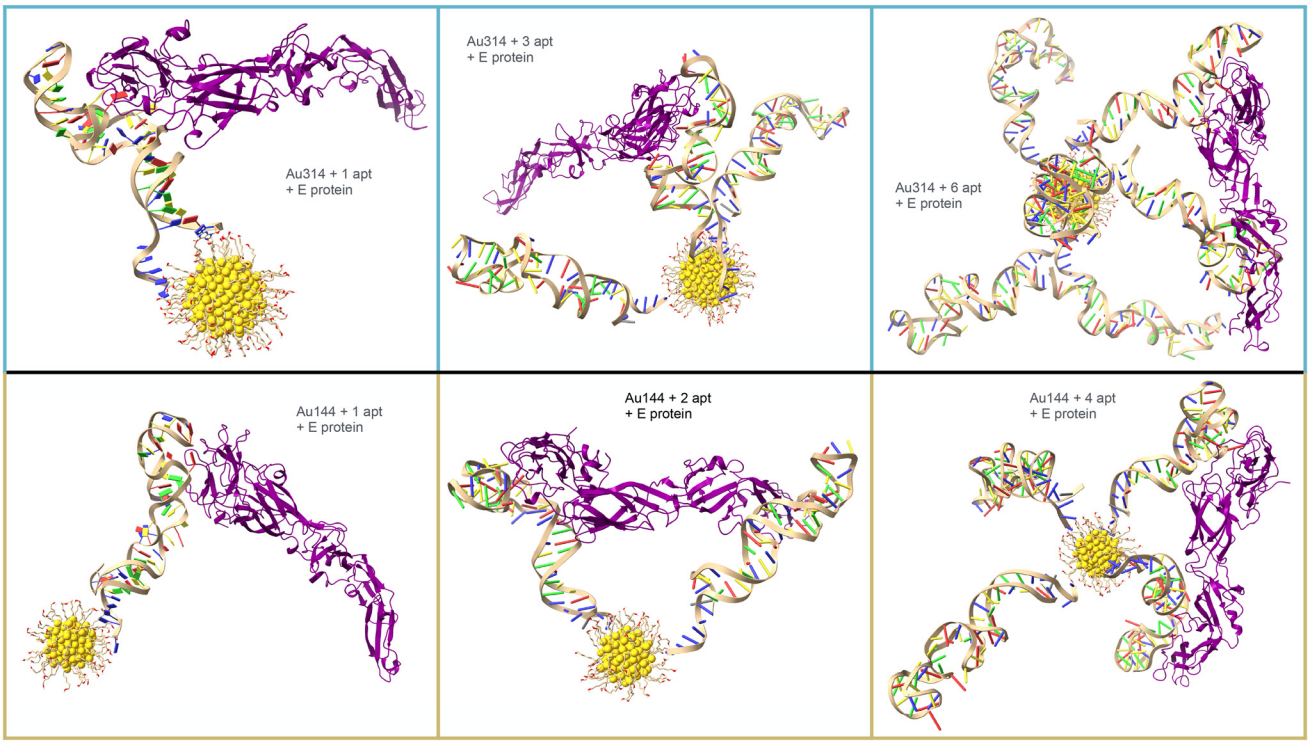


Fig. 3 Systems that have been simulated in the presence of the E protein. For each data set we show the most representative configuration among the final frames of 5 independent simulations.

ation and electrostatic potential (ESP) fitting, was based on procedures previously validated against experimental data.[3,27,48]

A detailed table containing the RESP charges computed by PyRED and the adjusted values we used in our simulations is available in section S2 of SI.

2.3 Molecular dynamics

Molecular dynamics simulations are primed and performed with GROMACS 2023.3.[52,53] The initial structures, prepared as explained in the previous section, are enclosed in a simulation box large enough to prevent the interaction of the system with its periodic images along the whole trajectory: this requires a cube with edge length ranging from 124 Å for the smallest system to 279 Å for systems with many aptamers and the E protein. The system is then solvated in a number of TIP3P[54] water molecules ranging from 60k to 700k according to the periodic cell size. The number of Na^+ and Cl^- ions needed to neutralize the system and to reach the physiological salt concentration of 150 mM l^{-1} is computed using the SLTCAP method.[55] Each aptamer carries a charge equal to $-59e$: it is convenient to initially place the counterions in proximity to the DNA backbone rather than randomly in the simulation box; this is accomplished by using the tleap program in the AmberTools24 suite.[56]

Velocities for all atoms are randomly generated according to a Maxwell distribution at temperature 300 K, then the system undergoes two rounds of equilibration with position restraints preventing the displacement of all heavy atoms of the solute: first, 100 ps of isothermal dynamics, under velocity-rescaling thermostat,[57] followed by other 100 ps of isothermal-isobaric dynamics with added Parrinello–Rahman pressure coupling at 1 bar.[58,59] Finally, position restraints are turned off and we perform 500 ns of isothermal-isobaric molecular dynamics, with standard leap-frog algorithm and 2 fs integration step. Each simulation is repeated 5 times, with different random seeds for the velocity generation. A snapshot of the system containing the position of each atom is saved every 0.1 ns; each MD trajectory comprises therefore 5000 frames.

3 Results and discussion

3.1 Radius of the NPs in this study

The gold NP cores used in this study have an approximately spherical shape. At various points of the data analysis reported in the following sections, we will need an estimate of the radius of this sphere. Here, we offer three such estimates, based on slightly different definitions, which will be used for different purposes.

3.1.1 Radius of the gold core. An estimate of the radius can be obtained by measuring the distance of each gold atom from the center of mass of the NP core. The resulting distribution (Fig. S2 in SI) shows very clearly the several shells that constitute the gold core, with the outermost shell comprised

of 48 and 30 atoms for $Au_{314}(SR)_{96}$ and $Au_{144}(SR)_{60}$, respectively. Focusing on this external shell of gold atoms we get a radius of (11.38 ± 0.13) Å for the larger NP, and (8.86 ± 0.14) Å for the smaller one, the uncertainty being the standard deviation (Fig. S3 in SI). Finally, if we add the van der Waals radius of gold (1.66 Å), they become R^{l}_{Au} = (13.04 ± 0.13) Å and R^{s}_{Au} = (10.52 ± 0.14) Å, respectively. These last two values are those that have been used to compute the aptamer density in Table 2.

3.1.2 Radius of the sulfur spherical shell. For the analysis of the ligand angles defined in section 3.4, it is more useful to define a spherical surface that approximates the position of the sulfur atoms belonging to the thiolate ligands. For each frame of each trajectory, the center and radius of such a sphere are determined according to a nonlinear least-squares fit performed with the Levenberg–Marquardt algorithm.[60] In this way, we find that the mean radius of $Au_{144}(SR)_{60}$ is R^{S}_{s} = (9.33 ± 0.05) Å, while the mean radius of $Au_{314}(SR)_{96}$ is R^{l}_{S} = (12.03 ± 0.07) Å (the uncertainty is the standard deviation of the values measured in different frames).

The very small variance of the radius of the sphere, with no discernible trend in its time evolution, is a confirmation that the force field for the gold core of the NP is very stable for the entire duration of the simulation. In order to quantify the quality of the spherical approximation, we can compute the Root Mean Square Deviation (RMSD) of the sulfur atoms from the optimized spherical surface. It turns out that the spherical approximation is excellent both for the smaller NP (average RMSD is 0.20 Å, only 2.2% of the radius) and for the larger NP (average RMSD is 0.38 Å, which represent a 3.2% of the radius).

3.1.3 Radius of the coated NP. The thiolate ligand coating is important for the drag force and therefore must be taken into account when computing the diffusion coefficient of the NP. This can be done by measuring the distance between the center of the gold core and the oxygen atom of each ligand (the same as used in section 3.4 to define the ligand angle ϑ, see Fig. S5 in SI). The resulting distributions are not Gaussian (nor are they expected to be), and are represented in Fig. S4 in SI. Using, as in the previous sections, the mean as the most representative value and the standard deviation as a measure of uncertainty, we get R^{l}_{c} = (18.5 ± 1.0) Å for $Au_{314}(SR)_{96}$ and R^{s}_{c} = (15.8 ± 1.2) Å for $Au_{144}(SR)_{60}$.

3.2 Configurations of the Tr-WNV-37 aptamer

Aptamers are *very* flexible. A standard RMSD-based clustering of the DNA backbone structures sampled during simulation of the aptamer-coated NPs shows that it is not possible to find a single reasonably representative structure: tentative clusters are scantly populated unless we choose a cutoff radius so large that the similarity between structures clustered together becomes vague.

Nonetheless, certain distinctive traits do emerge. Residues from 3 to 12 are complementary to residues from 55 to 46; this is somehow underestimated by the prediction software we used to create a starting configuration for the aptamer (see section 2.1.1). Even if they are not initially in contact (Fig. 1 (B)), during the simulation matching base-pairs tend to find each other on times of the order ~100 ns and form an almost full turn of double-helix. Once formed, this helix is generally conserved until the end of the MD run. By analogy with the nomenclature commonly used for DNA hairpins, we call this region the *stem*. In contrast, the central portion of the aptamer (residues from 13 to 54), exibit sparse and complementary pairing, a highly heterogeneous backbone configuration, and nucleobases that frequently protrude from the backbone, making them available for interactions with the surrounding environment. We refer to this region as the *bulge*: note that the nomenclature of secondary structure motifs in nucleic acid aptamers is not yet standardized.[42] A typical structure exemplifying these characteristics is represented in Fig. 4.

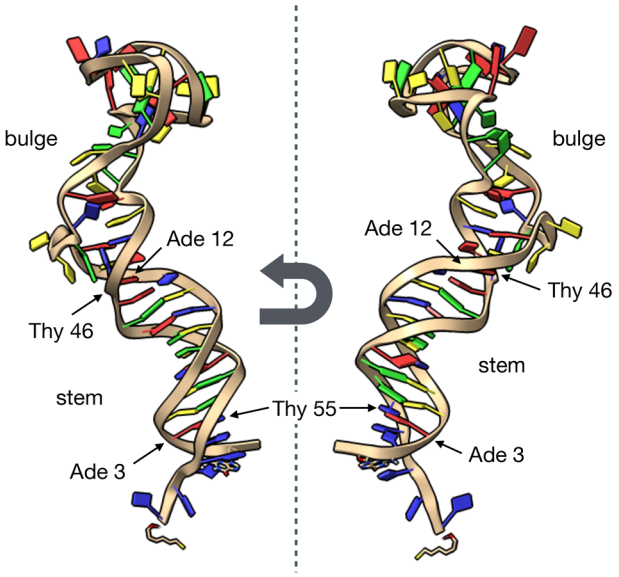


Fig. 4 One typical configuration of the aptamer Tr-WNV-37. We can distinguish a stem region of complementary base pairs (residues 3–-12 matching 46–55) and double-helix morphology from a bulge where complementary base-pairing is sparse and the backbone very flexible.

3.3 Diffusion coefficient

The variety of simulated systems (Fig. 2) makes it possible to investigate how the core size and the number of aptamers influence the diffusivity of the NP. The diffusion coefficients D estimated as explained in section S3 of SI are represented in Fig. 5, together with the theoretical predictions based on Stokes–Einstein law for the NPs without aptamers, which diffuse as spherical objects.

It should be kept in mind that although the maximum likelihood estimator used by us outperforms other competing methods, it is still affected by a rather large intrinsic uncertainty (its theoretical variance is 1/3 of the real value of D[61]), essentially due to the fact that we are estimating an ensemble average from a single—and not so long—trajectory. Moreover, the number of simulations in each data set is small (10 for the NPs without aptamers, 5 for the others). In this light,

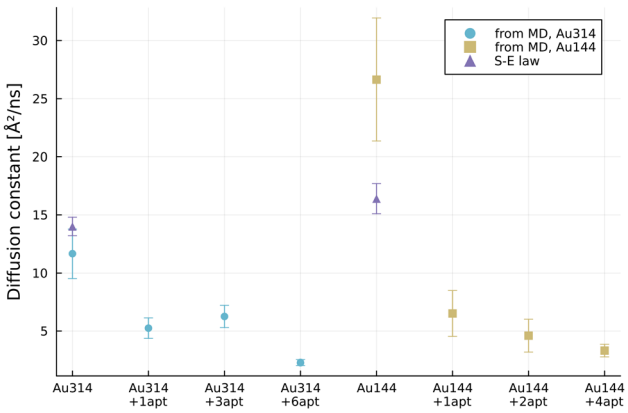


Fig. 5 Estimated diffusion coefficients D for different NPs. Blue circles and tan squares represent the median of the D values computed for each trajectory, with error bars signaling the standard error of the median as given by statistical bootstrap. Purple triangles are the predictions of Stokes–Einstein law for the NPs without aptamers.

anomalies like the unexpectedly large values of D for $Au_{144}(SR)_{60}$ and $Au_{314}(SR)_{96}$ with 3 aptamers should not be considered particularly significant.

However, despite these limitations, the correlation between the radius of gyration R_g and the hydrodynamic radius R_h (*i.e.*, the radius of a spherical particle which has the same diffusion coefficient as the functionalized NP) shown in Fig. 6 is reasonably linear, and the trend for smaller and larger NPs is very similar. This suggests a way to give a very rough estimate of R_h starting from a measure of R_g, which, unlike R_h, requires modest computational resources.

3.4 Ligand angles

In the initial configuration of the functionalized NP the thiolate ligands stick perpendicularly out of the surface, but within the first few tens of ps of MD simulation such artificial arrangement relaxes to a more physically realistic conformation. For each ligand, we call ϑ the angle between the vector going from the sulfur to the hydroxyl oxygen of the 6-mercapto-1-hexanol ligands and the normal to the NP surface at the position of the sulfur, as in Fig. S5 in SI. The normal vector to the surface is the vector from the center of the sphere, determined as explained in section 3.1.2, to the sulfur atom to which the ligand is attached.

In the analysis of the distribution of ligand angles, we have found it useful to distinguish between the ligands on $Au_{144}(SR)_{60}$ and the ligands on $Au_{314}(SR)_{96}$, and between ligands to which an aptamer is attached (which we indicate with the residue name ET6) and ligands with a free end (residue name ETO). The four resulting distributions are represented in Fig. 7. As shown in Table 3, the mean values are very similar and around 35°; the standard deviations are also quite similar and around 15° for all data sets. The main difference is found in the right tail of the distributions: of course, free ligands find it easier to reach large values of ϑ than ligands with an attached aptamer. Less obvious is the fact that ligands on the small NP have a significantly better chance of bending close to the surface than similar ligands on the large NP. The variation in ligand angle reflects a curvature-related effect, influencing local surface accessibility and ion interactions. While standard deviations are large, the trend is consistent across simulations, highlighting a systematic impact on nanoparticle stability and binding behavior relevant for design considerations.

3.5 Ion distribution

The equilibrium distribution of ion positions in relation to the NP is an important factor in the design of a biosensor, as plas-

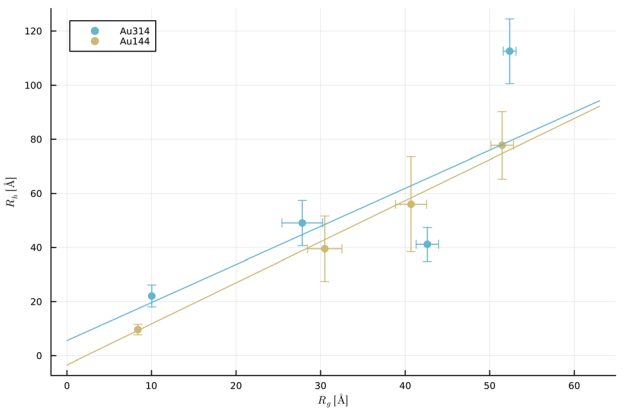


Fig. 6 Comparison between the hydrodynamic radius R_h and the radius of gyration R_g for the NPs of Fig. 2. Error bars on R_g are too small to be visible for the NPs without aptamers. The slope of the linear regression lines for larger and smaller NPs are similar: 1.4 and 1.5, respectively.

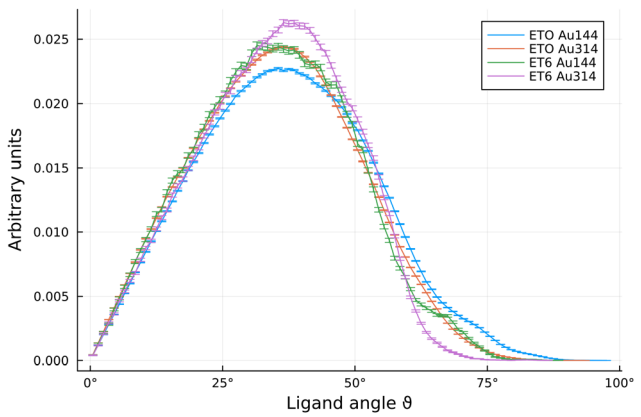


Fig. 7 Distribution of the ligand angle ϑ. Error bars are the square root of the number of observations in a bin. Thiolate residues with one free end (ETO) are distinguished from residues attached to an aptamer (ET6).

Table 3 Statistics about ligand angles

Data set	Mean	St. dev.	$p(\vartheta > 60°)$
ETO Au_{144}	37.0°	16.1°	0.080
ETO Au_{314}	35.2°	15.2°	0.053
ET6 Au_{144}	34.9°	14.9°	0.047
ET6 Au_{314}	34.9°	13.9°	0.021

monic NPs are very sensitive to any chemical change occurring in their environment. In this section we are interested in the distribution of the minimum distance d between each ion and any part of the functionalized NP (including the gold core, its thiolate coating and the aptamers). As mentioned in section 2.3, when the system is prepared for simulation, 59 cations per aptamer are initially placed in the vicinity of the negatively charged DNA backbone. During the equilibration phase and the production run, however, both the counterions added to neutralize the system and the ions added to mimic physiological conditions are free to move. It turns out that for all our systems the ion distance distribution converges to equilibrium in less than 50 ns (see section S4 in SI): this allows us to average over all subsequent time snapshots, thus achieving the high resolution distributions showed in this section.

3.5.1 Nanoparticle alone. In Fig. 8 we observe that in the region $d > 25$ Å the cation and the anion distributions are indistinguishable. In the short-distance region, the increased statistics makes it possible to clearly distinguish three peaks located at 2.4 Å, 4.2 Å, and 6.7 Å. The distributions obtained from the other data sets exhibit the same features (see Fig. S9 in SI). So the cations placed to neutralize the system stay close to the NP during the entire simulation, while the others are as randomly distributed as the anions; but where around the NP do they spend more time?

To answer this question, we restrict our analysis to the cations of the highest peak, those that establish contact with the NP, which we define by the condition $d < 3.0$ Å. For each cation in contact with the NP, we record the closest residue; the resulting histogram is represented in Fig. 9.

We can clearly distinguish three regions: residues 1–144 are gold atoms of the NP core, 145–202 are ligands of the NP coating, 203–263 and 264–324 are two aptamers. The rare cations that reach the gold core tend to stay there for extended periods, kinetically trapped by the ligands surrounding them, hence leaving a sizable trace in the histogram of contacts. In comparison, more cations establish a contact with one of the thiolate ligands, but such contacts tend to be shorter-lived. All other data sets exhibit the same features (see Fig. S10 in SI).

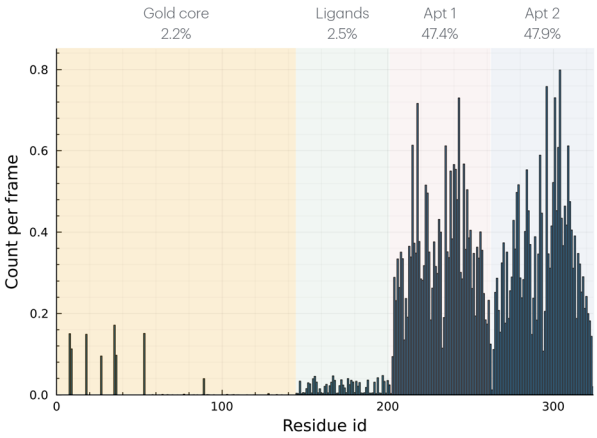


Fig. 9 Histogram of cation contacts by closest NP residue. Data obtained by summing all frames with $t > 50$ ns from 5 independent runs of NP Au_{144} with 2 aptamers. Residues 1–144 are gold atoms forming the NP core, 145–202 are thiolate ligands, 203–263 and 264–324 are the two aptamers. The percentages at the top of the figure break down the total cation contacts by group of residues.

Remarkably, the pattern of contacts established with each aptamer appears to be repeated. This suggests the idea of focusing on the cations in contact with an aptamer, thus observing the relative propensity of different aptamer residues to be in contact with positive ions. Where more than one aptamer is present, this is achieved by classifying contacts by aptamer residue (instead of by NP residue, as in Fig. 9), thus summing together the histogram portions that refer to all aptamers.

In order to be able to compare different data sets, it is necessary to normalize the respective histograms. If we divide the tally by the number of frames multiplied by the number of aptamers, we measure the frequency of ion contacts as a function of the aptamer residue. Note that, as a single aptamer residue can happen to be the closest part of the NP for more than one ion at the same time, here the word "frequency" should not be intended as a quantity necessarily less than one. The plots thus obtained for each individual data set are available as Fig. S11 in SI, while the cumulative histogram that takes into account all data sets is represented in Fig. 10.

It is clear that certain residues are systematically more capable than others of retaining positive ions in their proximity, and this effect does not depend on the nucleotide base (see Fig. S12 in SI). An attempt to link such propensity to an easily computed property such as the solvent-accessible surface area (SASA, computed using the algorithm introduced by Eisenhaber *et al.*[62] as implemented in the GROMACS library of analysis tools) yields very poor correlation coefficients, as shown in Fig. S13 in SI.

3.5.2 In the presence of the protein. The same analysis of ion positions with respect to the NP can be repeated for simulations in the presence of the E protein and is shown in detail

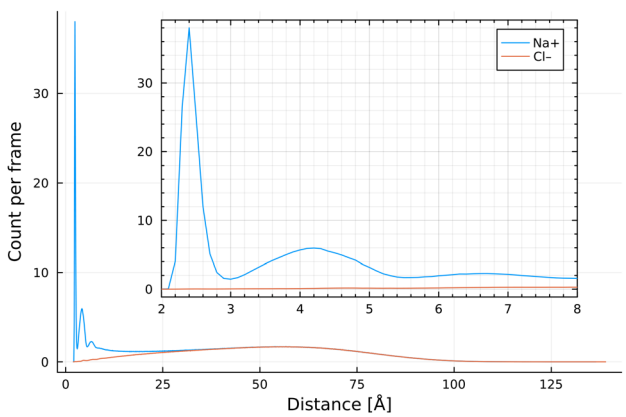


Fig. 8 Distribution of ion distances to the NP Au_{314} with 6 aptamers obtained by summing all frames with $t > 50$ ns from 5 independent trajectories. The inset is a magnified detail of the same graph.

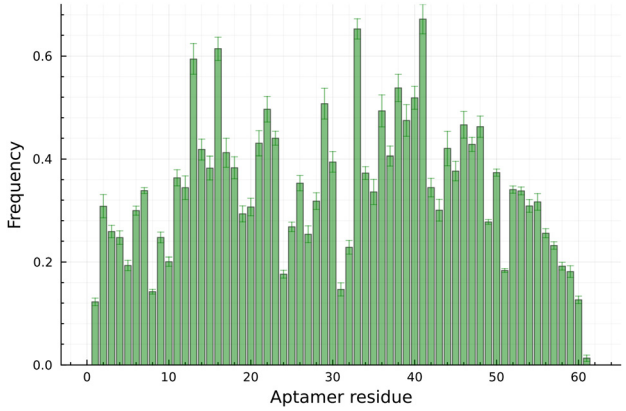


Fig. 10 Frequency of cation contacts by aptamer residue. The height and the uncertainty of each bar are the mean and the standard error (*i.e.* standard deviation over the square root of sample size), respectively, of the heights of the same bar across all data sets (individually represented in Fig. S11 in SI).

in section S5 in SI. More interesting, however, is the effect of the presence of the protein on the frequency of cation contacts by aptamer residue, the quantity shown for the case without protein in Fig. 10. It seems natural to expect that the aptamer residues directly involved in the interaction with the protein should lose some of their availability to establish contacts with cations, and this is indeed the case at least for some of them. As discussed in section 3.6, the aptamer residue by far most engaged in interactions with the protein is Ade 38 (see Fig. 16); if we measure its cation contact frequency across all MD traces and compare the results with and without protein we obtain Fig. 11, where the effect of the presence of the protein is clearly in accordance with our expectation.

However, if we extend this analysis to the other key aptamer residues in Fig. 16, the situation is not so clear (see Fig. S15 in SI). Instead of focusing on the cation contacts of a single aptamer residue, it is more informative to look at the global modifications undergone by the histogram in Fig. 10 when the protein is present, as illustrated in Fig. 12.

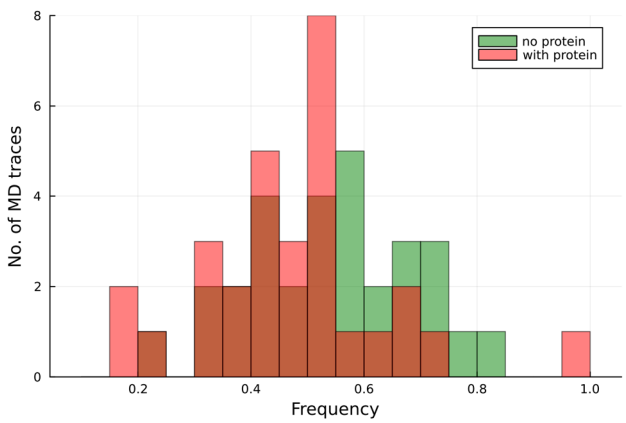


Fig. 11 Frequency of cation contacts with the aptamer residue Ade 38 computed in all MD trajectories.

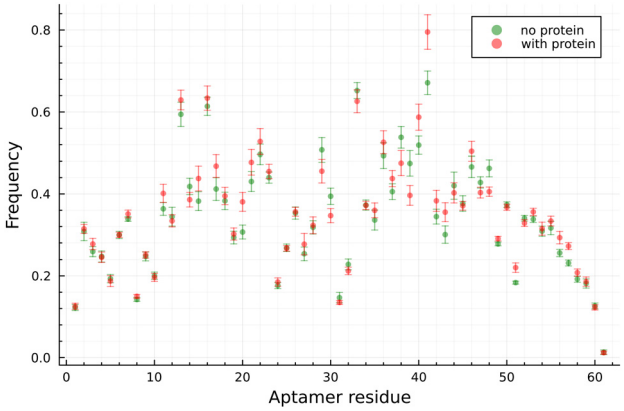


Fig. 12 Frequency of cation contacts by aptamer residue, as in Fig. 10. Effect of the presence of the protein.

It can be seen that the presence of the protein does not really affect the frequency of cation contacts for the residues in the lower stem of the aptamer (residue numbers less than 10 or more than 50). For the central residues, the protein induces a certain displacement of the cations in contact with the aptamer, but this effect tends to be localized in a small region: as shown in Fig. 13, the probability of establishing contact with a cation lost by Ade 38 and Ade 39 is gained by the neighboring residues.

In summary, our ion distribution analysis reveals that positively charged ions initially placed near the aptamer backbone remain stably associated with the NP system throughout the simulations. While ions at greater distances exhibit a random, bulk-like distribution, a significant population of cations remains tightly localized near the NP surface, forming distinct peaks at short distances. These tightly bound cations primarily associate with the gold core and thiolate ligands, with some exhibiting long-lived interactions indicative of kinetic trapping.

Remarkably, specific aptamer residues consistently attract cations, highlighting a non-uniform charge distribution along

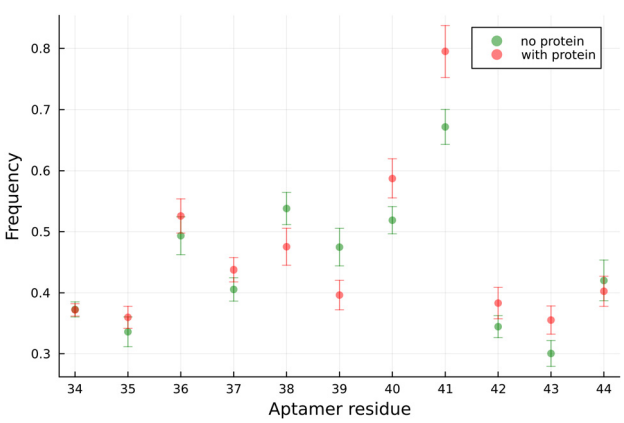


Fig. 13 Detail of Fig. 12, showing the region around the residue Ade 38, which is the most engaged in interactions with the E protein.

the DNA strand. This behavior is not correlated with solvent-accessible surface area (SASA), suggesting that local electrostatic environments and aptamer conformation play a more dominant role in ion association.

When the E protein of West Nile virus is present, the ion distribution remains mostly unchanged in the bulk region, but slight shifts in local cation binding are observed. Notably, residues involved in protein binding, particularly Ade 38 and Ade 39, exhibit reduced cation interactions, in line with their involvement in protein contacts. The displaced cation interactions are redistributed to nearby residues, suggesting a localized reorganization of the electrostatic environment upon protein binding.

Overall, these findings emphasize the dynamic and selective nature of cation association in aptamer-functionalized nanoclusters and underscore its relevance for biosensor performance, where surface charge modulation can significantly influence both target binding and signal readout.

3.6 Contacts between aptamer and protein

We define a contact between a residue in the protein and a residue in the functionalized NP if the minimum distance between any pair of atoms in the residues is less than 3.0 Å. The first thing to observe about these contacts is that their time evolution is a dynamic process (see Fig. S16–S22 in SI). Individual contacts can be established in the course of the simulation, then interrupted for some time, and then appear again. Changes in the relative position of protein and NP may result in the sudden breakdown of a whole set of contacts that is swiftly replaced by a different one. A contact can be ever present in one MD trajectory, and completely absent in another one. The variability among independent runs of the same system does not appear to be any less than the variability among different data sets. Yet, when we look at the whole body of data from our simulations, some clear indications emerge.

In the following, we report statistics about the frequency of contacts by protein residue, aptamer residue, and specific pair of protein-aptamer residues. We will call *primary* contacts those formed by the aptamer originally bound to the protein (to which we will refer as the *first* aptamer), and *secondary* contacts those established in the course of the simulation by other aptamers (if present). Such secondary contacts are by no means a rare occurrence in our simulations: in fact, the mean number of secondary contacts per frame (computed over all traces where more than one aptamer is present) is 1.03, which suggests that finding other aptamers in interaction with the E protein is more likely than not.

3.6.1 Contacts by protein residue. The average number of primary contacts per protein residue has been computed pooling the observations in each frame of each MD trajectory where the protein is present, for a total of 175 000 time snapshots. The residues with a mean number of contacts larger than 0.5 are listed in Table 4.

The residues most likely to be in contact with the first aptamer belong to domain DIII of the protein (residues from 298 to 403, see Fig. 1(A)), with the exception of a few residues in DI, as shown in Fig. 14. The residue ARG 388 stands out as by far the most important for interactions with the first aptamer; its very exposed position enables it to regularly engage in more than one contact simultaneously.

Table 4 Mean number of primary contacts between the E protein and the first aptamer

Type	Id	Mean	Type	Id	Mean
ARG	388	2.571	PRO	351	0.859
VAL	304	1.186	LYS	337	0.747
LEU	349	1.068	SER	16	0.723
LYS	36	1.003	PRO	339	0.697
THR	350	0.929	GLY	387	0.609
SER	306	0.900	GLU	390	0.570
LYS	307	0.881	SER	341	0.533

Secondary contacts, on the other hand, involve almost exclusively the domain DII, and no individual residue stands out as particularly more attractive than others (see Table 5). We can notice, however, a prevalence of secondary contacts with the positively charged (ARG, LYS) or the polar (SER, THR, ASN, GLN) amino acids, a feature that is not unexpected due to the negatively charged nature of the aptamer. As DII is normally connected to the virus envelope, it is not as available *in vivo* to interact with aptamers as is in our simulations. However, we can expect secondary contacts to be common and

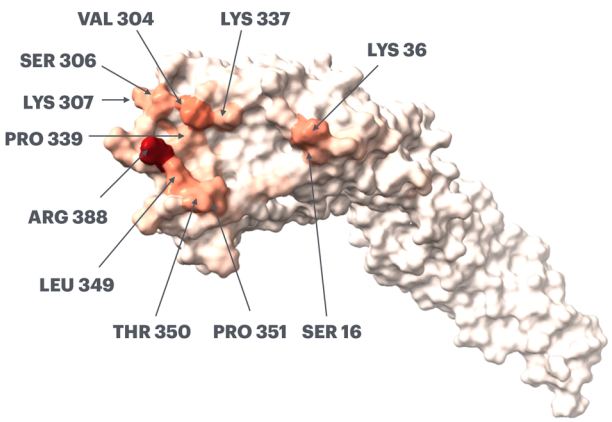


Fig. 14 Reference structure of the E protein color-coded by average number of primary contacts with the aptamer.

Table 5 Mean number of secondary contacts between the E protein and other aptamers

Type	Id	Mean	Type	Id	Mean	Type	Id	Mean
ASN	199	3.2%	GLN	94	2.6%	VAL	96	1.8%
LYS	110	3.1%	ARG	234	2.5%	SER	175	1.7%
ARG	193	3.0%	THR	177	2.4%	SER	230	1.6%
LYS	249	2.9%	ARG	93	2.2%	TYR	176	1.6%
ARG	215	2.6%	LYS	132	2.1%	TRP	225	1.6%
THR	231	2.6%	ARG	85	2.0%	ARG	57	1.6%

to play a role in *in vitro* experiments where the E protein is detached by the rest of the virus.

3.6.2 Contacts by aptamer residue. The aptamer residues for which we observe a mean number of primary contacts larger than 0.5 are reported in Table 6.

As ARG 388 was pre-eminent among the protein residues active in forming contacts, so is Ade 38 among the aptamer residues. The fact that in most of the time snapshots this adenine sticks clearly out of the DNA backbone allows it to establish contacts with as much as 8 amino acids at the same time (see Fig. S18 in SI, the panel about run 2), thus creating a very strong binding to the protein. The other top residues belong mostly in the bulge, with the exception of a group of them at around half of the aptamer length, as shown in Fig. 15.

Residues involved in establishing secondary contacts are considerably less localized, as they can be found along the entire length of the aptamer (see Table 7).

3.6.3 Contacts by specific pair of residues. After what we have seen in the two previous sections, it does not come as a surprise that in all data sets, the most frequent interaction is between ARG 388 and Ade 38, which is present in 84.3% of the 175 000 frames analyzed from all trajectories. If we consider a subset of the contact map containing the most frequently observed primary contacts (Fig. 16), it is evident the key role played by nucleotide Ade 38.

Table 6 Mean number of primary contacts between the first aptamer and the E protein

Type	Id	Mean	Type	Id	Mean
Ade	38	5.470	Cyt	10	0.864
Ade	29	2.192	Cyt	15	0.795
Gua	39	1.304	Gua	30	0.791
Ade	14	1.098	Ade	28	0.584
Ade	37	1.066	Thy	22	0.547
Cyt	21	0.969	Thy	11	0.533

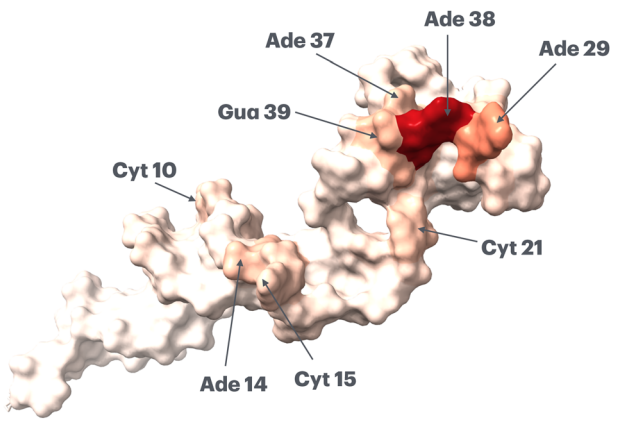


Fig. 15 Reference structure of an aptamer color-coded by average number of primary contacts with the E protein.

Table 7 Mean number of secondary contacts between other aptamers and the E protein

Type	Id	Mean	Type	Id	Mean	Type	Id	Mean
Cyt	35	5.1%	Cyt	20	3.8%	Gua	39	3.1%
Gua	1	4.9%	Ade	14	3.7%	Gua	50	3.0%
Ade	29	4.4%	Cyt	51	3.6%	Cyt	21	2.7%
Cyt	15	4.0%	Thy	27	3.6%	Ade	3	2.6%
Ade	49	3.9%	Ade	38	3.4%	Cyt	16	2.5%

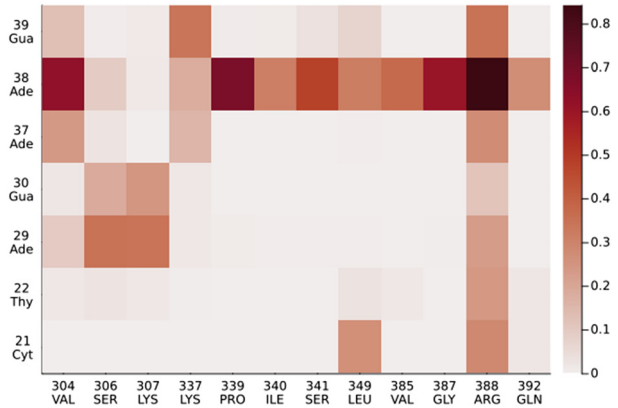


Fig. 16 A portion of the contact map showing the 19 most frequent contacts between protein and aptamer.

Apart from the contacts established by the aptamer originally bound to the protein, in the simulations of the NPs with more than one aptamer it is quite common to observe the formation of other, non-specific, contacts. A portion of the contact map large enough to contain the 15 most frequent non-specific interactions is shown in Fig. 17.

3.7 Free energy of interaction

We estimate the free energy of interaction between the E protein and the functionalized NP by using the Molecular

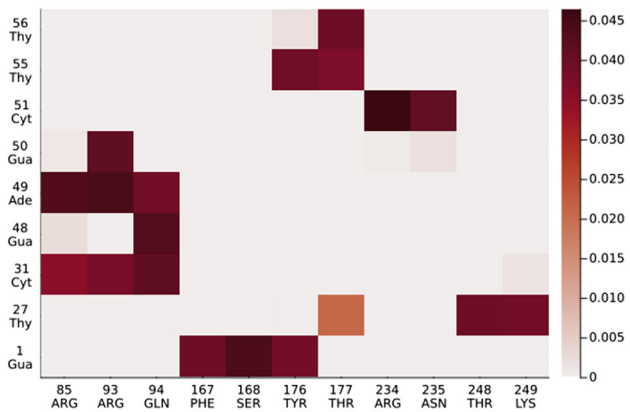


Fig. 17 A portion of the contact map restricted to non-specific interactions between the protein and aptamers not originally bound to it, showing the 15 most frequent contacts.

Mechanics Generalized Born Surface Area (MMGBSA) method[63] as implemented in the gmx_MMPBSA software.[64,65] For each MD trajectory, we select 100 snapshots from the last 100 ns and compute the ΔG on each frame separately. The mean value and standard deviation of the estimated ΔG for each trajectory are represented in Fig. 18, and are color-coded according to the mean number of secondary contacts per frame.

Comparison with the data set labeled apt where the aptamer is not attached to any NP shows that immobilization does not spoil the interaction with the protein. If anything, it slightly improves the binding; in Fig. 19 we compare two distributions: one represents the values of ΔG computed in all trajectories belonging to the apt data set, where the aptamer is free. The other is the distribution of ΔG values obtained in all trajectories where the aptamer is immobilized on a NP, and no secondary contacts are observed (the white dots of Fig. 18).

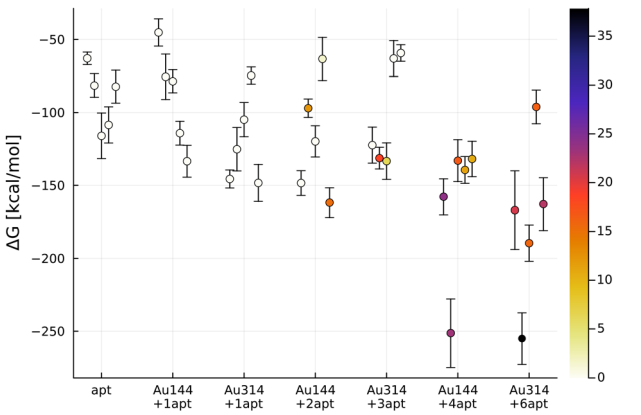


Fig. 18 Free energy of interaction between the E protein and the functionalized NP as evaluated with the MMGBSA method. All independent runs are showed as distinct points, error bars are given by the standard deviation. The color scale refers to the mean number of secondary contacts.

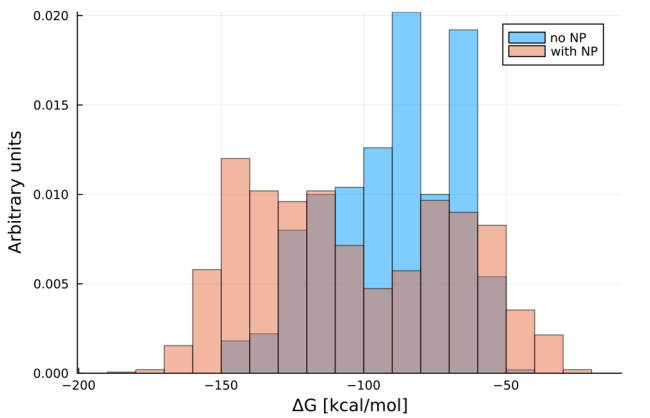


Fig. 19 Free energy of interaction between the E protein and the functionalized NP as evaluated with the MMGBSA method. Comparison between the data set where the aptamer is free, and the data sets where the aptamer is immobilized on a NP, but only primary interactions are observed.

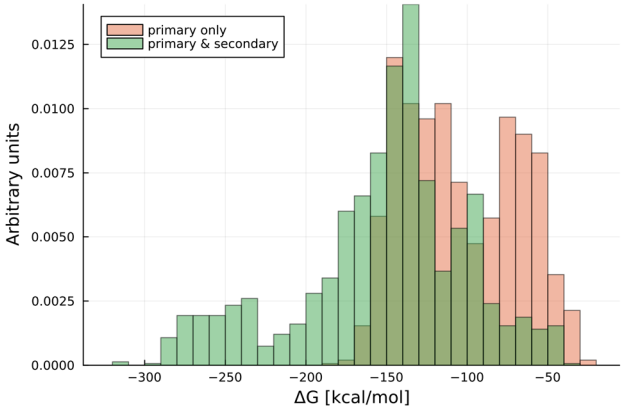


Fig. 20 Free energy of interaction between the E protein and the functionalized NP as evaluated with the MMGBSA method. Comparison between the trajectories where only primary interactions are observed and those where secondary contacts are present as well.

The mean values are −90 kcal mol^{-1} and −104 kcal mol^{-1}, respectively, but the clear bimodality of the second distribution suggests that further distinctions could be pursued in this case.

On the other hand, it is very clear that the presence of secondary contacts strengthens the interaction between the functionalized NP and the protein, as shown in Fig. 20. That this effect is actually due to secondary contacts is confirmed by the Pearson correlation coefficient −0.74 between the estimated ΔG and the mean number of secondary contacts per frame (the color scale in Fig. 18).

4. Summary and conclusions

In this study, we employed atomistic molecular dynamics simulations to investigate the structural and dynamic behavior of gold nanoclusters functionalized with the experimentally validated Tr-WNV-37 aptamer, a molecular recognition element specific for the West Nile virus (WNV) envelope (E) protein. By analyzing two representative gold nanoparticle (AuNP) sizes, namely $Au_{344}(SR)_{96}$ and $Au_{144}(SR)_{60}$, we accurately estimated nanoparticle radii, aptamer packing density, and ligand organization, key parameters for evaluating biosensor architecture and responsiveness. Diffusion and ion distribution analyses revealed that both nanoparticle size and aptamer loading markedly influence particle mobility and local electrostatic organization, with cations showing adaptive rearrangements upon protein binding. These observations underline the importance of ionic screening and nanoparticle surface chemistry in maintaining aptamer–protein recognition under physiological, high-salt conditions that often hinder biosensor performance.

At the molecular level, our simulations demonstrate that the Tr-WNV-37 aptamer retains strong and specific interactions with the E protein upon immobilization, primarily through residue Ade 38 engaging ARG 388 in domain III, with

secondary contacts involving domain II. This multi-point binding mechanism suggests a stable and cooperative recognition mode, which could enhance target capture efficiency and selectivity in practical sensing configurations.

The development of *ad hoc* force-field parameters for aptamer-functionalized AuNPs ensured accurate representation of hybrid bio-nano interfaces, establishing a transferable computational framework for future studies. Together, these results identify nanoparticle size, aptamer surface density, and ionic environment as key design variables for tuning sensor performance and reliability.

This work represents a critical step toward designing an aptamer-based diagnostic platform for West Nile virus (WNV) detection, to be validated *via* Surface Plasmon Resonance (SPR) measurements. Recent studies, such as those by Petronella *et al.*,[66,67] have shown the sensitivity of plasmonic nanostructures for label-free pathogen detection. These molecular-level details, difficult to access experimentally, will be directly compared with forthcoming SPR data to ensure strong correspondence between theory and experiment. Ongoing work focuses on deriving coarse-grained (CG) models from our atomistic simulations, following established methodologies for protein–nanoparticle systems,[68,69] to investigate aptamer–E protein recognition and selectivity under realistic, crowded *in vitro* conditions.

Overall, our work demonstrates that gold nanocluster size, aptamer flexibility and surface coverage all play critical roles in modulating interaction dynamics with protein targets. The results highlight key design parameters for optimizing aptamer-functionalized nanomaterials in biosensing platforms. This integrated computational approach provides a powerful framework for predicting and rationalizing aptamer-target interactions and can be extended to other viral or protein biomarkers to guide the development of next-generation diagnostic nanodevices.

Author contributions

A. Mossa MD investigation, analysis of MD data, writing of the initial draft. G. Brancolini conceptualization, supervision, writing of the initial draft and funding acquisition.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the supplementary information (SI). Supplementary Information reports the protein and aptamer sequences of the simulated constructs, derivation of partial charges, the estimation of diffusion coefficients and analysis of ion distributions around the studied systems. See DOI: https://doi.org/10.1039/d5nr03228h.

Acknowledgements

G. B. and A. M. gratefully acknowledge G. Schifino and S. Corni for their valuable discussions and collaborative contributions to this work. G. B. acknowledges the support by EU within the NextGenerationEU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. PE00000007, INF-ACT). We acknowledge the CINECA award under the ISCRA initiative, for the availability of high-performance computing resources and support.

References

1 L. Barzon, M. Pacenti, S. Ulbert and G. Palù, *Expert Rev. Anti-Infect. Ther.*, 2015, **13**, 327–342.
2 H. Nwanosike, F. M. Green, K. O. Murray, J. E. Weatherhead and S. E. Ronca, in *West Nile Virus: Methods and Protocols*, ed. F. Bai, Springer US, New York, NY, 2023, pp. 119–125.
3 M. Ray, G. Brancolini, D. C. Luther, Z. Jiang, R. Cao-Milán, A. M. Cuadros, A. Burden, V. Clark, S. Rana, R. Mout, R. F. Landis, S. Corni and V. M. Rotello, *Nanoscale*, 2022, **14**, 2411–2418.
4 M. Ruan, M. Seydou, V. Noel, B. Piro, F. Maurel and F. Barbault, *J. Phys. Chem. B*, 2017, **121**, 4071–4080.
5 P. Si, N. Razmi, O. Nur, S. Solanki, C. M. Pandey, R. K. Gupta, B. D. Malhotra, M. Willander and A. de la Zerda, *Nanoscale Adv.*, 2021, **3**, 2679–2698.
6 P. Mertinková, E. Mochnáčová, K. Bhide, A. Kulkarni, Z. Tkáčová, J. Hruškovicová and M. Bhide, *Sci. Rep.*, 2021, **11**, 20131.
7 I. M. Khoris, F. Nasrin, A. D. Chowdhury and E. Y. Park, *Anal. Chim. Acta*, 2022, **1207**, 339817.
8 E. Iswardy, T.-C. Tsai, I.-F. Cheng, T.-C. Ho, G. C. Perng and H.-C. Chang, *Biosens. Bioelectron.*, 2017, **95**, 174–180.
9 P. R. Abraham, R. Bharathy, N. Pradeep Kumar and A. Kumar, *Sci. Rep.*, 2021, **11**, 23699.
10 A. Vázquez-Guardado, F. Mehta, B. Jimenez, A. Biswas, K. Ray, A. Baksh, S. Lee, N. Saraf, S. Seal and D. Chanda, *Nano Lett.*, 2021, **21**, 7505–7511.
11 X.-X. Ding, X.-F. Li, Y.-Q. Deng, Y.-H. Guo, W. Hao, X.-Y. Che, C.-F. Qin and N. Fu, *PLoS One*, 2014, **9**, e108623.
12 A. Krüger, A. P. de Jesus Santos, V. de Sá, H. Ulrich and C. Wrenger, *Pharmaceuticals*, 2021, **14**, 622.
13 Y. Aslan, M. Atabay, H. K. Chowdhury, I. Göktürk, Y. Saylan and F. Inci, *Biosensors*, 2023, **13**, 569.
14 K. H. Lee and H. Zeng, *Anal. Chem.*, 2017, **89**, 12743–12748.
15 K.-I. Matsunaga, M. Kimoto, V. W. Lim, H. P. Tan, Y. Q. Wong, W. Sun, S. Vasoo, Y. S. Leo and I. Hirao, *Nucleic Acids Res.*, 2021, **49**, 11407–11424.

16 R. Thevendran, S. Rogini, G. Leighton, A. Mutombwera, S. Shigdar, T.-H. Tang and M. Citartan, *Biology*, 2023, **12**, 722.
17 L. M. d. Morais, T. S. Chaves, M. A. Medeiros, K. A. B. Pereira, P. B. Jurgilas, S. M. Barbosa de Lima, S. Missailidis and A. M. Bispo de Filippis, *Viruses*, 2022, **14**, 1867.
18 C. Stephens, N. M. Goodey and U. Gubler, *Anal. Biochem.*, 2025, **703**, 115890.
19 Z. Chen, L. Hu, B.-T. Zhang, A. Lu, Y. Wang, Y. Yu and G. Zhang, *Int. J. Mol. Sci.*, 2021, **22**, 3605.
20 S. J. Lee, J. Cho, B.-H. Lee, D. Hwang and J.-W. Park, *Biomedicines*, 2023, **11**, 356.
21 A. A. Buglak, A. V. Samokhvalov, A. V. Zherdev and B. B. Dzantiev, *Int. J. Mol. Sci.*, 2020, **21**, 8420.
22 S. Dutta, S. Corni and G. Brancolini, *Int. J. Mol. Sci.*, 2022, **23**, 1484.
23 J. Song, Y. Zheng, M. Huang, L. Wu, W. Wang, Z. Zhu, Y. Song and C. Yang, *Anal. Chem.*, 2020, **92**, 3307–3314.
24 K. L. Rhinehardt, G. Srinivas and R. V. Mohan, *J. Phys. Chem. B*, 2015, **119**, 6571–6583.
25 W. Phanchai, U. Srikulwong, A. Chompoosor, C. Sakonsinsiri and T. Puangmali, *Langmuir*, 2018, **34**, 6161–6169.
26 H. Park, N. Kwon, G. Park, M. Jang, Y. Kwon, Y. Yoon, J. An, J. Min and T. Lee, *Bioelectrochemistry*, 2023, **154**, 108540.
27 G. Brancolini, V. M. Rotello and S. Corni, *Int. J. Mol. Sci.*, 2022, **23**, 2368.
28 G. E. Nybakken, C. A. Nelson, B. R. Chen, M. S. Diamond and D. H. Fremont, *J. Virol.*, 2006, **80**, 11467–11474.
29 S. Mukhopadhyay, B.-S. Kim, P. R. Chipman, M. G. Rossmann and R. J. Kuhn, *Science*, 2003, **302**, 248.
30 E. C. Meng, T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris and T. E. Ferrin, *Protein Sci.*, 2023, **32**, e4792.
31 R. Anandakrishnan, B. Aguilar and A. V. Onufriev, *Nucleic Acids Res.*, 2012, **40**, W537–W541.
32 O. Lopez-Acevedo, J. Akola, R. L. Whetten, H. Grönbeck and H. Häkkinen, *J. Phys. Chem. C*, 2009, **113**, 5035–5038.
33 S. Malola, L. Lehtovaara, J. Enkovaara and H. Häkkinen, *ACS Nano*, 2013, **7**, 10263–10270.
34 S. Franco-Ulloa, L. Riccardi, F. Rimembrana, M. Pini and M. De Vivo, *J. Chem. Theory Comput.*, 2019, **15**, 2022–2032.
35 S. Giberti, S. Dutta, S. Corni, M. Frasconi and G. Brancolini, *Nanoscale*, 2025, **17**, 4389–4399.
36 G. Brancolini, L. Bellucci, M. C. Maschio, R. Di Felice and S. Corni, *Curr. Opin. Colloid Interface Sci.*, 2019, **41**, 86–94.
37 A. Imanparast, N. Attaran, H. Eshghi and A. Sazgarnia, *Iran. J. Basic Med. Sci.*, 2022, **25**, 970–979.
38 M. V. Cherrier, B. Kaufmann, G. E. Nybakken, S. Lok, J. T. Warren, B. R. Chen, C. A. Nelson, V. A. Kostyuchenko, H. A. Holdaway, P. R. Chipman, R. J. Kuhn, M. S. Diamond, M. G. Rossmann and D. H. Fremont, *EMBO J.*, 2009, **28**, 3269–3276.
39 N. R. Markham and M. Zuker, in *Bioinformatics: Structure, Function and Applications*, ed. J. M. Keith, Humana Press, Totowa, NJ, 2008, pp. 3–31.
40 M. Popenda, M. Szachniuk, M. Antczak, K. J. Purzycka, P. Lukasiak, N. Bartol, J. Blazewicz and R. W. Adamiak, *Nucleic Acids Res.*, 2012, **40**, e112.
41 J. Sarzynska, M. Popenda, M. Antczak and M. Szachniuk, *Proteins: Struct., Funct., Bioinf.*, 2023, **91**, 1790–1799.
42 S. Ochoa and V. T. Milam, *ACS Synth. Biol.*, 2025, **14**, 3049–3064.
43 J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytč, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis and J. M. Jumper, *Nature*, 2024, **630**, 493–500.
44 M. Pons, M. Perenon, H. Bonnet, E. Gillon, C. Vallée, L. Coche-Guérente, E. Defrancq, N. Spinelli, A. van der Heyden and J. Dejeu, *Analyst*, 2022, **147**, 4197–4205.
45 G. C. P. van Zundert, J. P. G. L. M. Rodrigues, M. Trellet, C. Schmitz, P. L. Kastritis, E. Karaca, A. S. J. Melquiond, M. van Dijk, S. J. de Vries and A. M. J. J. Bonvin, *J. Mol. Biol.*, 2016, **428**, 720–725.
46 M. van Dijk, A. D. J. van Dijk, V. Hsu, R. Boelens and A. M. J. J. Bonvin, *Nucleic Acids Res.*, 2006, **34**, 3317–3325.
47 G. Brancolini, D. Toroz and S. Corni, *Nanoscale*, 2014, **6**, 7903–7911.
48 S. Dutta, S. Corni and G. Brancolini, *Int. J. Mol. Sci.*, 2021, **22**, 3624.
49 V. Lindahl, A. Villa and B. Hess, *PLoS Comput. Biol.*, 2017, **13**, e1005463.
50 I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpí, C. González, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case and M. Orozco, *Nat. Methods*, 2016, **13**, 55–58.
51 E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak and F.-Y. Dupradeau, *Nucleic Acids Res.*, 2011, **39**, W511–W517.
52 D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.
53 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.
54 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
55 J. D. Schmit, N. L. Kariyawasam, V. Needham and P. E. Smith, *J. Chem. Theory Comput.*, 2018, **14**, 1823–1827.
56 D. A. Case, H. M. Aktulga, K. Belfon, D. S. Cerutti, G. A. Cisneros, V. W. D. Cruzeiro, N. Forouzesh, T. J. Giese, A. W. Götz, H. Gohlke, S. Izadi, K. Kasavajhala, M. C. Kaymak, E. King, T. Kurtzman, T.-S. Lee, P. Li, J. Liu,

T. Luchko, R. Luo, M. Manathunga, M. R. Machado, H. M. Nguyen, K. A. O'Hearn, A. V. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, A. Risheh, S. Schott-Verdugo, A. Shajan, J. Swails, J. Wang, H. Wei, X. Wu, Y. Wu, S. Zhang, S. Zhao, Q. Zhu, T. E. Cheatham III, D. R. Roe, A. Roitberg, C. Simmerling, D. M. York, M. C. Nagan and K. M. Merz Jr, *J. Chem. Inf. Model.*, 2023, **63**, 6183–6191.

57 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

58 M. Parrinello and A. Rahman, *Phys. Rev. Lett.*, 1980, **45**, 1196–1199.

59 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

60 K. Levenberg, *Q. Appl. Math.*, 1944, **2**, 164–168.

61 D. Boyer, D. S. Dean, C. Mejía-Monasterio and G. Oshanin, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2012, **85**, 031136.

62 F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander and M. Scharf, *J. Comput. Chem.*, 1995, **16**, 273–284.

63 I. Massova and P. A. Kollman, *Perspect. Drug Discovery Des.*, 2000, **18**, 113–135.

64 B. R. Miller III, T. D. McGee Jr., J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, *J. Chem. Theory Comput.*, 2012, **8**, 3314–3321.

65 M. S. Valdés-Tresanco, M. E. Valdés-Tresanco, P. A. Valiente and E. Moreno, *J. Chem. Theory Comput.*, 2021, **17**, 6281–6291.

66 F. Petronella, D. Stoia, Y. Ziai, F. Zaccagnini, V. Scognamiglio, D. Maniu, C. Rinoldi, M. Focsan, A. Antonacci, F. Pierini and L. De Sio, in *Novel Optical Materials*, ed. I. C. Khoo, F. Simoni and C. Umeton, World Scientific, 2023, pp. 155–194.

67 F. Petronella, D. De Biase, C. Santini, A. Avitabile, M. L. Sforza, F. Zaccagnini, A. d'Alessandro and L. De Sio, *Mater. Adv.*, 2025, **6**, 1107–1118.

68 G. Brancolini, H. Lopez, S. Corni and V. Tozzini, *Int. J. Mol. Sci.*, 2019, **20**, 3866.

69 G. Brancolini and V. Tozzini, *Front. Mol. Biosci.*, 2019, **6**, 50.