Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Quantitative threshold modelling of targeting-ligand shielding by protein corona on engineered extracellular vesicles: an analytic coverage threshold verified by Monte-Carlo simulation

Zhuofan Shen

bbv215290@uic.es

Universitat Internacional de Catalunya https://orcid.org/0009-0005-4304-8391

Research Article

Keywords: extracellular vesicles, exosome engineering, protein corona, steric shielding, ligand density, surface coverage, Monte-Carlo model

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11804242/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Targeting ligands displayed on engineered extracellular vesicles (EVs) lose function once a plasma protein corona adsorbs onto the same membrane. The field has generally responded by increasing ligand density, on the working assumption that a high enough display density must eventually out-compete steric occlusion. This assumption is examined quantitatively here.

A parameterised Monte-Carlo model was built on the EV surface as a sphere, with ligands and adsorbed corona proteins represented as spherical caps of specified projected area and with overlap tested by exact cap geometry. Four ligand classes (RVG peptide, 2.2 nm²; nanobody, 7.5 nm²; transferrin, 14.0 nm²; IgG, 32.0 nm²), five EV diameters (40–160 nm), ligand densities from 0.1 to 10 per 100 nm² and nominal corona coverages from 0.05 to 0.70 were evaluated. Two accessibility metrics were defined: a binary metric, in which a ligand counts only if no adsorbed protein touches it, and a fractional metric, in which the ligand surface left uncovered is measured directly by surface sampling.

The central result is negative with respect to density. Under the binary metric the accessible fraction is independent of ligand density: across a hundred-fold density range the absolute deviation was at most 0.031, and no systematic trend was present. Increasing ligand density therefore cannot recover accessibility, because shielding is governed by the number of adsorbed proteins per unit area, not by the number of ligands competing with them. Density buys additional accessible copies at a fixed exchange rate, not a higher fraction.

Accessibility is instead set by a single dimensionless group, and the model admits a closed form: ALF = exp[-f (1 + (S_ligand/S_corona)^(1/2))²], where f is the nominal corona coverage and the two S terms are projected areas. Fitted to 280 simulated means the exponents returned 1.000, 1.000 and 2.002 with a root-mean-square deviation of 0.0083, so the relation is essentially exact rather than approximate. It is scale-free between 40 and 160 nm to within 8%, and the resulting half-accessibility coverages are 0.34, 0.22, 0.16 and 0.10 for the four ligand classes, compared with the 0.50 typically assumed under human plasma exposure. Expressed in realised rather than nominal occupancy, which differ by a factor of 1.27 at that coverage, the IgG threshold requires less than 10% of the membrane to be occupied by protein.

The practical conclusion is that corona suppression, not ligand over-display, is the operative design lever, and that ligand molecular size sets how much corona a formulation can tolerate. The results delimit one mechanism under stated idealisations and provide design references rather than in-vivo predictions; all sampling rules, parameters, verification criteria and an algorithmic specification are reported so that the numbers can be regenerated from the text.

1. Introduction

Extracellular vesicles are endogenous nanoscale carriers with a phospholipid membrane, a protein and RNA cargo and, in principle, no requirement for synthetic scaffolding. Their biodistribution in vivo is governed chiefly by cell source, route of administration and the surface targeting moieties they carry [1],

and this has made surface engineering the practical route to selective delivery. The recurring difficulty is that targeting behaves differently in serum-free binding assays and in animals, and the in-vitro-to-in-vivo drop-off remains a central obstacle to translation.

Quantification of the ligand layer itself has improved markedly. Single-particle nano-flow cytometry now resolves conjugation ratio, ligand density and the uniformity of surface distribution, and comparative work across lipid, protein and membrane-insertion chemistries reports conjugation efficiencies above 90% with ligand densities of several to tens of ligands per 100 nm² under optimised conditions [2]. Density is therefore no longer an unmeasured variable, and the standard engineering response to a targeting failure has become to raise it: more ligands per particle, on the reasoning that a denser display should survive occlusion better.

A major contributor to the loss is protein corona formation. Upon contact with plasma, abundant circulating proteins adsorb within minutes onto the EV surface, and mass-spectrometry profiling of corona-coated EVs identifies a reproducible set of shared corona proteins dominated by apolipoproteins, complement components and fibrinogen [3]. Crucially, corona formation is not a passive decoration: it changes the effective biological identity of the particle, and for EVs it has been shown to redirect cellular uptake and circulation behaviour in a composition-dependent manner [5, 6]. Corona formation has also been examined directly in the context of EV surface engineering, where chemisorbed and physisorbed antibody presentations converge to comparable behaviour once a biomolecular corona is present [4].

The mechanism most directly relevant to targeting is steric shielding: adsorbed plasma proteins physically occupy membrane area and cover surface-displayed ligands, blocking receptor engagement. This phenomenon is well documented for synthetic nanoparticles, where corona composition predicts cellular interaction more accurately than particle size, aggregation state or surface charge [7], and where the corona is understood to supply the biological identity perceived by cells [8]. For EVs the same physics applies, but the consequence for ligand display has not been converted into a design rule. In particular, it has not been established whether the accessible fraction can be raised at all by increasing ligand density, because simple geometric reasoning suggests that it cannot: if a ligand is lost whenever an adsorbed protein cap touches it, then the probability of being touched depends on the number of proteins per unit area and on the two cap sizes, and not on how many other ligands share the surface. That reasoning is tested here rather than assumed.

Three quantitative questions follow. First, is ligand density actually a control variable for accessibility, or is the field optimising the wrong parameter? Second, if accessibility is instead governed by corona coverage, how much coverage can each ligand class tolerate, and does the answer depend on vesicle diameter? Third, can the answer be delivered in a form compact enough to use during formulation design, with a stated error and a stated sensitivity to the input parameters? Existing EV pharmacokinetic and biodistribution models treat targeting efficiency as a fixed input and do not resolve corona-dependent ligand loss, so none of them addresses these questions.

This work answers them with a deliberately narrow model. The scope is restricted to geometric steric shielding; ligand proteolysis, EV aggregation, mononuclear phagocyte clearance and endosomal-lysosomal processing are excluded. A Monte-Carlo framework evaluates the accessible-ligand fraction over the design space, is reduced to critical coverage thresholds, and admits a closed form whose exponents are recovered from the simulation rather than imposed. The contributions are: (i) a demonstration that ligand density is not a threshold variable for accessibility, which redirects the design question; (ii) an analytic coverage threshold, ALF = exp[-f(1+(S_ligand/S_corona)^(1/2))²], whose exponents are verified against 280 simulated means with a root-mean-square deviation of 0.0083; and (iii) a quantified distinction between nominal coverage, the quantity usually specified, and realised area occupancy, together with a first-order sensitivity band on every reported threshold. All procedures, parameters and evaluation criteria are contained in this manuscript.

2. Materials and Methods

2.1 Model physical assumptions

1. Small extracellular vesicles are approximated as smooth spheres of homogeneous surface chemistry.

2. Targeting ligands are presented radially outward and each occupies a defined circular membrane-projection area derived from its molecular dimensions; the projection is treated as a spherical cap rather than a planar disc.

3. Corona-forming plasma proteins are modelled as spherical adsorbates distributed randomly and independently over the EV surface. A ligand is classified as shielded if its cap overlaps the cap of any adsorbed protein, and as accessible otherwise.

4. Corona coverage denotes the fraction of total EV surface area nominally assigned to projected protein caps, obtained as the number of proteins multiplied by their projection area divided by the vesicle area. Because caps overlap one another at high coverage, the realised occupied area is smaller than this nominal value, and both quantities are reported separately in Section 3.6.

5. Ligand-ligand steric repulsion, ligand clustering and cooperative ordering are not represented; baseline ligand placement is spatially uniform and random. Ligand density is additionally constrained by a packing ceiling, since the projected ligand area cannot exceed the available membrane area; the ceiling is set at 0.85 of the surface.
6. All physicochemical inputs are taken from previously published experimental literature; no parameter is introduced to improve fit agreement.
7. Two accessibility metrics are defined and reported separately. The binary accessible-ligand fraction (ALF) is the proportion of displayed ligands whose caps are touched by no adsorbed protein. The fractional accessibility (ALF_area) is the fraction of the nominal ligand footprint that remains uncovered, measured by direct surface sampling; it therefore counts a partially occluded ligand as partially

accessible. The binary metric answers whether a ligand can engage a receptor at all, the fractional metric answers how much epitope surface survives.

2.2 Input parameter space

Five EV hydrodynamic diameters were simulated: 40, 70, 100, 130 and 160 nm, spanning the small-EV window. Ligand surface densities of 0.1, 0.2, 0.4, 0.8, 1.5, 3.0, 5.0, 7.0 and 10.0 ligands per 100 nm² were evaluated, truncated where the packing ceiling is reached: the ceiling corresponds to 38.6, 11.3, 6.07 and 2.66 ligands per 100 nm² for RVG peptide, nanobody, transferrin and IgG respectively. The lower part of this range brackets the several-to-tens of ligands per 100 nm² reported for optimised post-production conjugation chemistries [2]; the upper part is included to test whether accessibility can be rescued by crowding ligands together.

Four ligand classes were used, each with a literature-derived membrane projection area: RVG-derived short peptide, 2.2 nm²; single-domain nanobody, 7.5 nm²; transferrin glycoprotein, 14.0 nm²; and full-length IgG antibody, 32.0 nm². These classes represent the range of molecular size encountered in practice, from a short receptor-binding peptide to an intact immunoglobulin. Their projection areas relative to the corona protein give the ratios S_ligand/S_corona = 0.18, 0.63, 1.17 and 2.67, which is the quantity that the analytic result below turns out to depend on.

Nominal corona surface-coverage fractions of 0.05 to 0.70 were simulated in fourteen steps, with 0.5 taken as the reference human-plasma condition. An average corona-protein projection area of 12.0 nm² was assigned, consistent with the abundant small plasma proteins, principally albumin, that dominate early corona formation [3, 4].

2.3 Computational implementation

Monte-Carlo surface sampling was implemented in Python 3 with NumPy and SciPy, using a fixed random seed. For each combination of EV diameter, ligand density, ligand class and corona coverage the following procedure was applied:

1. Ligand centres were generated as uniform random points on the sphere surface by sampling the height coordinate uniformly on [-1, 1] and the azimuthal angle uniformly on [0, 2π). Sampling the polar angle uniformly instead would concentrate points at the poles and bias coverage, so the height-uniform construction is used throughout.

2. Corona-protein centres were generated by the same uniform sampling procedure, with their number set by round(fA/S_corona), where $A = 4\pi R^2$ is the vesicle area. Ligand copy numbers were set by round(ρA/100).

3. Each molecule's projected area was converted to the half-angle of its spherical cap, $\theta = \arccos(1 - S/2\pi R^2)$. A ligand was counted as shielded when the great-circle separation between its centre and any protein centre fell below the sum of the two half-angles, which is equivalent to testing overlap of the

caps themselves. The test was implemented as a chord-length neighbour query, $|p_i - p_j| \le 2 \sin((\theta_i + \theta_j)/2)$, using a k-d tree.

4. The binary accessible fraction was computed as the number of unshielded ligands divided by the total displayed. Fifteen independent replicates were run for each coverage-sweep condition and twenty-five for each density-sweep condition; mean and standard deviation were recorded.

5. The fractional accessibility was estimated by sampling a large number of uniform surface points per replicate and classifying each point as covered by at least one ligand cap, by at least one protein cap, both or neither. The accessible ligand area was taken as the fraction of points in the ligand-covered set that lie outside every protein cap, multiplied by the vesicle area, and divided by the nominal ligand area $n_{ligand} \cdot S_{ligand}$. Sixty thousand points were used per replicate in the density sweep, giving a sampling standard error below 0.002 on a coverage estimate near 0.5.

Downstream reduction was performed on the tabulated means rather than on individual replicates. Critical coverages were obtained by log-linear interpolation of mean ALF against nominal coverage for each combination of ligand class and diameter, and inverted at ALF = 0.50 and ALF = 0.10. A closed form was then fitted to all 280 tabulated means by non-linear least squares from multiple starting values, and the root-mean-square deviation between fit and table is reported as the fit error. Because the fit is derived from the same sweep used for the interpolation, the two descriptions are not independent evidence; the fit serves to compress the table and to expose the functional form, and the deviation between fitted and interpolated thresholds is reported explicitly.

All simulation settings, sampling rules, overlap criteria and fitted parameters are documented in this manuscript, so that the workflow can be reproduced from the text alone; no external repository is required. Appendix A restates the procedure in algorithmic form and lists the verification checks applied to the output.

2.4 Sensitivity analysis design

Because every input parameter carries measurement uncertainty, the sensitivity of the reported thresholds was evaluated by re-running the coverage sweep with perturbed inputs, not by propagation through the fit. Three perturbations were applied individually:

1. Ligand projection area varied by ± 20%, reflecting the range obtained when a ligand is approximated by different conformations or hydration states.
2. Average corona-protein projection area varied between 10 and 14 nm² around the assigned 12 nm², reflecting the difference between a small-protein-dominated and a mixed-composition corona.
3. A ± 10% uncertainty in the measured value of corona coverage itself, which does not move the threshold but moves the margin between the formulation and the threshold; it is reported as a fractional margin rather than as a threshold change.

The interval reported in Section 3.7 is the outer envelope of the first two perturbations and should be read as the practical uncertainty band on a threshold value rather than as a confidence interval.

2.5 Model boundary statement

The model addresses steric overlap-driven ligand shielding only. It does not represent ligand proteolysis, EV aggregation, altered membrane fluidity, ligand clustering, exchange between hard and soft corona components, inter-individual variation in plasma composition, ligand orientation, or downstream clearance and intracellular processing. The thresholds are therefore design references derived under stated idealisations, and they do not substitute for experimental measurement of ligand accessibility in serum or in vivo.

3. Results

3.1 Zero-coverage baseline and behaviour of the two metrics

With zero adsorbed protein no overlap is possible and the computed ALF is 1.000 for every EV diameter, ligand class and density tested (0.1 to 3.0 ligands per 100 nm²), with no drift attributable to cap arithmetic. This is the expected numerical counterpart of a serum-free binding assay and confirms that all shielding reported below originates from the simulated corona.

At reference coverage (nominal 0.50) and a reference density of 1.5 ligands per 100 nm², the binary accessible fraction on a 100 nm vesicle is 0.363, 0.200, 0.116 and 0.029 for RVG peptide, nanobody, transferrin and IgG respectively, each the mean of fifteen replicates. The fractional accessibility under the same conditions is 0.605, 0.568, 0.553 and 0.537. The two metrics therefore differ by a factor of 1.7 to 18 depending on ligand size, which is not a numerical artefact but a statement about what is being asked: at half coverage, roughly one ligand in three is still wholly untouched if the ligand is a short peptide, whereas for an intact antibody almost every molecule is touched, yet more than half of the antibody’s projected surface remains uncovered. A formulation judged by epitope exposure and a formulation judged by intact-ligand count are being assessed on different scales, and the two scales are reported separately throughout.

3.2 Density invariance of the binary metric

The decisive test is whether increasing ligand density raises accessibility. For a 100 nm vesicle at nominal coverage 0.50, mean binary accessibility was measured at nine densities spanning a hundred-fold range. The absolute variation across that range was 0.031 for RVG peptide, 0.016 for nanobody, 0.013 for transferrin and 0.005 for IgG, with no systematic trend in either direction:

- - RVG peptide: 0.368, 0.338, 0.369, 0.358, 0.363, 0.360, 0.364, 0.359, 0.365 at 0.1, 0.2, 0.4, 0.8, 1.5, 3.0, 5.0, 7.0 and 10.0 ligands per 100 nm² (corona proteins 1309, ligands 31 to 3142).
- - Nanobody: 0.213, 0.213, 0.198, 0.201, 0.200, 0.203, 0.204, 0.197, 0.205.

- - Transferrin: 0.129, 0.118, 0.123, 0.119, 0.116, 0.117, 0.116 at 0.1 to 5.0 ligands per 100 nm², the range being truncated at the packing ceiling.
- - IgG: 0.031, 0.031, 0.034, 0.029, 0.030 at 0.1 to 1.5 ligands per 100 nm², truncation reflecting the 2.66-per-100-nm² packing ceiling.

The deviations are of the order of the replicate standard error and show a flat rather than a rising profile: the largest departure, 0.023 at 0.2 ligands per 100 nm² for the peptide, is 1.7 standard errors from the mean of that series, and no departure at one density repeats at the next. The result follows directly from the geometry: a ligand is lost when a protein cap reaches it, and the rate at which caps reach a given point is fixed by the number of proteins per unit area; adding more ligands adds more targets but does not reduce the per-target hit rate. What density does change is the absolute number of accessible copies, which rises linearly: on a 100 nm vesicle at coverage 0.50, RVG displayed copies increase from 31 to 3142 across the density range and accessible copies from 11 to 1148, whereas IgG accessible copies rise only from 1 to 14 as display rises from 31 to 471. Density therefore purchases copies at a fixed exchange rate, and the exchange rate, not the density, is what a formulation can move.

At 40 nm the same sweep is no longer clean, because the copy numbers are small: at 0.1 ligands per 100 nm² a 40 nm vesicle carries 5 ligands and at 0.2 it carries 10, so the accessible fraction can only take values in steps of 0.2 and 0.1. Deviations of up to 57% in the 40 nm IgG case correspond to differences of 0.01 in absolute fraction and reflect integer counting rather than a physical density effect. Where the absolute copy number is below about fifty, the binary metric should be read as quantised.

3.3 Coverage sweep and critical coverage thresholds

Because density is not a control variable, the sweep was re-parameterised over nominal coverage. Binary accessibility falls monotonically with coverage for every ligand class, and the rate of fall is set by ligand size. On a 100 nm vesicle at 1.5 ligands per 100 nm², the measured values at nominal coverages of 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60 and 0.70 were:

- - RVG peptide: 0.904, 0.819, 0.669, 0.552, 0.445, 0.363, 0.294, 0.245.
- - Nanobody: 0.850, 0.725, 0.535, 0.376, 0.272, 0.202, 0.151, 0.106.
- - Transferrin: 0.810, 0.646, 0.428, 0.264, 0.181, 0.115, 0.073, 0.047.
- - IgG: 0.712, 0.512, 0.257, 0.126, 0.060, 0.029, 0.015, 0.009.

Inverting these series gives the critical coverage at which half the displayed ligands remain untouched. For 100 nm vesicles the values are 0.346, 0.219, 0.157 and 0.103 for RVG peptide, nanobody, transferrin and IgG respectively, each interpolated from fifteen replicates per point. The corresponding 10%-accessibility coverages are reachable within the simulated range only for the two larger ligands: 0.537 for transferrin and 0.341 for IgG. For the peptide and the nanobody the accessibility at the highest simulated coverage (0.70) is still 0.245 and 0.106, so a 10% target would require occupancy beyond 70% of the membrane surface, which is beyond both the simulated range and the regime in which the geometric approximation remains sensible.

Vesicle diameter changes these thresholds much less than the field's emphasis on size optimisation would suggest. Across 40 to 160 nm the half-accessibility coverage ranged over 0.336–0.363 for the peptide, 0.206–0.219 for the nanobody, 0.157–0.160 for transferrin and 0.099–0.103 for IgG, that is a relative spread of 7.8, 5.9, 1.7 and 4.6%. Diameter therefore does not provide an independent lever: a small vesicle and a large one of the same surface chemistry tolerate essentially the same occupancy fraction, and the small vesicle simply holds fewer copies of everything, which is why its absolute count is quantised as described in Section 3.2. This is a direct consequence of the geometry discussed next.

3.4 Closed form, and why it is scale-free

The exponent governing shielding can be written analytically. A ligand centre is hit when a corona centre falls inside an interaction cap whose half-angle is θ_ligand + θ_corona. The solid-angle fraction of that cap is (1 - cos(θ_ligand + θ_corona))/2, which for the small angles involved reduces to (θ_ligand + θ_corona)²/4. With n_corona = fA/S_corona proteins and A = 4πR², and using θ ≈ (S/πR²)^(1/2), the expected number of hits per ligand is:

E = f (1 + (S_ligand/S_corona)^(1/2))²

The vesicle radius cancels identically: the number of proteins grows as R² while the angular size of each cap shrinks as R⁻². The exponent is therefore a pure function of coverage and of the area ratio, and accessibility follows the Poisson form ALF = exp(-E):

ALF = exp[-f (1 + (S_ligand/S_corona)^(1/2))²]

Fitting this functional form with three free exponents, ALF = exp[-k f^p (1 + r)^q] with r = (S_ligand/S_corona)^(1/2), to all 280 simulated means returned k = 1.000, p = 1.000 and q = 2.002, with a root-mean-square deviation of 0.0083 and a largest single deviation of 0.035. The relation is thus recovered rather than imposed, and the fit cannot be dismissed as curve fitting: the exponents sit on their analytic values and the residual is at the level of replicate noise.

The law reproduces the sweep closely. For a 100 nm vesicle the analytic values at coverage 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60 and 0.70 are 0.903, 0.815, 0.665, 0.542, 0.442, 0.360, 0.294 and 0.240 for the peptide, and 0.707, 0.500, 0.250, 0.125, 0.063, 0.031, 0.016 and 0.008 for IgG; these differ from the measured values in Section 3.3 by at most 0.010 absolute for the peptide and 0.012 for IgG. The half-accessibility thresholds implied by the analytic form are 0.340, 0.216, 0.160 and 0.100 for the four classes, which agree with the values interpolated from the simulation to within 1.2% when averaged over the five diameters and to within 3.2% for any single diameter.

Two consequences follow. First, the threshold has a closed form that needs no computation: the coverage that preserves half the ligands is ln2/(1 + (S_ligand/S_corona)^(1/2))², which gives 0.34 for a peptide, 0.22 for a nanobody, 0.16 for transferrin and 0.10 for an antibody, and the coverage that preserves a tenth is ln10/(1 + r)². Second, the scale-free structure explains the small size dependence observed in Section 3.3: the only departures from exact scale invariance are the integer rounding of copy

numbers and, for the largest ligand on the smallest vesicle, the breakdown of the small-angle approximation, where the IgG value at 40 nm sits 3% below its large-vesicle value.

3.5 Fractional accessibility, crowding and accessible copy number

The binary metric is insensitive to density because it counts a ligand as lost the moment any protein touches it. The fractional metric behaves differently, since overlapping ligands consume each other’s footprint. At nominal coverage 0.50 on a 100 nm vesicle, fractional accessibility fell monotonically from the lowest density to the packing ceiling:

- - RVG peptide: 0.611, 0.580, 0.599, 0.599, 0.605, 0.586, 0.577, 0.558, 0.545 from 0.1 to 10.0 ligands per 100 nm², a fall of 0.066.
- - Nanobody: 0.599 at 0.1 and 0.429 at 10.0, a fall of 0.170 (28%).
- - Transferrin: 0.599 at 0.1 and 0.435 at 5.0, a fall of 0.164 (27%).
- - IgG: 0.618 at 0.1 and 0.480 at 1.5, a fall of 0.138 (22%).

The magnitude of the crowding penalty tracks the ligand footprint, because a large ligand at high density occupies area already occupied by its neighbours. At the packing ceiling the nominal ligand coverage of the membrane is 0.85 by construction, and the accessible fraction can no longer be raised by adding more ligand: the membrane has no room. This is the quantitative form of the intuition that over-conjugation eventually backfires, and it is a penalty on the fractional metric of up to 28%, not on the binary one.

The number that a delivery programme actually needs is neither fraction alone. It is the count of sterically available ligands per particle, which is the product of copy number and binary accessibility, and which scales linearly with density at fixed coverage. For a 100 nm vesicle at coverage 0.50 the accessible count is 11 at 0.1 ligands per 100 nm², 171 at 1.5 and 1148 at 10.0 for the peptide; 7 at 0.1, 94 at 1.5 and 644 at 10.0 for the nanobody; and only 4 at 0.1, 55 at 1.5 and 182 at 5.0 for transferrin, with IgG reaching 1 at 0.1 and 14 at 1.5 ligands per 100 nm². A formulation whose avidity requirement is a few dozen accessible ligands is therefore reachable with a peptide or a nanobody at modest density, and is reachable with IgG only near its packing ceiling and only if the coverage is brought below the threshold of Section 3.4.

3.6 Nominal versus realised coverage

The coverage parameter used above is nominal: it is the sum of projected protein areas divided by the vesicle area. Adsorbed caps overlap one another, so the area actually occupied is smaller. Direct surface sampling of the corona layer at nominal coverages of 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.50, 0.60 and 0.70 returned realised occupancies of 0.049, 0.095, 0.139, 0.180, 0.221, 0.259, 0.297, 0.330, 0.393, 0.455 and 0.504, that is ratios of 0.985, 0.945, 0.926, 0.900, 0.883, 0.862, 0.848, 0.825, 0.786,

0.759 and 0.721. The discrepancy grows with coverage and reaches 28% at the highest value. It is independent of vesicle diameter, differing by less than 0.006 between 40, 100 and 160 nm vesicles.

This matters because the two quantities are not interchangeable when a threshold is quoted against experimental data. The half-accessibility thresholds expressed in realised occupancy are 0.293, 0.195, 0.145 and 0.097 for the four ligand classes, against nominal values of 0.346, 0.219, 0.157 and 0.103. An experimentalist who quantifies corona loading from adsorbed protein mass, or who reports the fraction of surface area covered by imaging, is measuring the realised quantity, and a comparison made against the nominal threshold overstates the formulation's margin by roughly a quarter at the reference condition. Reporting the convention explicitly, as is done here, removes that source of ambiguity.

3.7 Sensitivity of the thresholds to input uncertainty

Re-running the coverage sweep with perturbed inputs moved every threshold by less than 17%, and never changed the ordering of the four ligand classes. Relative to the unperturbed half-accessibility coverage, the outer envelope over a ± 20% change in ligand projected area and a 10–14 nm² range in corona-protein area was − 7.3% to + 5.9% for the peptide, -8.6% to + 14.0% for the nanobody, -11.9% to + 13.1% for transferrin and − 10.4% to + 16.3% for IgG. The asymmetric direction is expected: reducing the corona-protein area at fixed coverage raises the protein count and therefore lowers the tolerated coverage, while enlarging the ligand footprint raises it.

In absolute terms the perturbed thresholds span 0.32–0.36 for the peptide, 0.20–0.24 for the nanobody, 0.14–0.18 for transferrin and 0.087–0.113 for IgG. Separately, a ± 10% error in the measured value of corona coverage does not move the threshold at all; it moves the margin between the formulation and the threshold by the same ± 10% relative amount, which is the more useful way to express it when the comparison, rather than the threshold, is the uncertain step.

3.8 Qualitative literature check

Reported ligand densities for RVG-modified and transferrin-modified EVs were placed into the reference condition (100 nm vesicle, nominal coverage 0.50) to obtain a predicted accessibility for published formulations. Reported densities of several ligands per 100 nm² [2] fall above the coverage-limited ceiling in the sense that they cannot raise the fraction: the predicted binary accessibility for a transferrin-displayed EV at 0.5 coverage is 0.116 whatever the density. Published formulations that retain targeting in plasma are therefore better explained by a corona that is lighter than the reference assumption, or by a display chemistry that presents the ligand away from the membrane, than by a high ligand count.

This step is a consistency check on plausibility, not a validation, and it cannot establish that steric shielding is the operative cause in any individual study, because those studies differ simultaneously in EV source, purity, dose, route and host species, and because several of them do not report coverage at all. Its only function is to confirm that the thresholds produced by the model fall in the range where the experimental literature reports both success and failure.

4. Discussion

The result that overturns the working assumption of the field is a negative one: ligand density is not a threshold variable for steric accessibility. Under an exact geometric model in which a ligand is lost when an adsorbed protein cap touches it, the accessible fraction varied by at most 0.031 across a hundred-fold density range on a 100 nm vesicle and showed no trend. The mechanism is straightforward once stated. Shielding is a race between adsorbed proteins and the membrane area they cover; the number of ligands sharing that area is irrelevant to the per-ligand hit rate, because ligands do not protect one another from proteins. A denser display therefore produces more accessible ligands in absolute terms, at the same fractional rate, and no amount of over-conjugation changes the fraction.

This reframes the design problem. If fraction cannot be bought with density, it must be bought by reducing occupancy, and the thresholds for that are now explicit: preserve half the ligands at nominal coverage below 0.34 for a peptide, 0.22 for a nanobody, 0.16 for transferrin and 0.10 for an antibody. For a 100 nm vesicle these correspond to 906, 573, 411 and 269 adsorbed proteins of 12 nm² projection area. Since plasma at the reference condition adsorbs roughly 1300 such proteins, the antibody threshold asks for a fivefold reduction in adsorbed protein number, which is a statement about how strong a corona-mitigation strategy has to be rather than about how many ligands to conjugate. The corresponding message for ligand selection is that molecular size is a design decision with a quantitative price: moving from a nanobody to an intact IgG roughly halves the tolerated coverage, 0.22 to 0.10, for the same nominal epitope specificity.

The closed form deserves emphasis because it is not an empirical fit wearing an equation's clothes. The exponent f(1 + (S_ligand/S_corona)^(1/2))² follows from the solid angle of the interaction cap and the R²-versus-R⁻² cancellation, and the three fitted exponents came back as 1.000, 1.000 and 2.002 against 280 means with a residual of 0.0083. A reader can therefore use the relation without trusting the simulation, and can check it independently against any single measurement of coverage and ligand size. The scale-free structure is a second dividend: diameter does not enter the threshold, which removes a parameter that the EV literature routinely treats as central to targeting performance, and identifies where diameter does still matter, namely in the absolute copy number on small vesicles and in the quantisation of the accessible fraction when that copy number is small.

Three practical implications follow directly:

1. Corona suppression, not ligand over-display, is the operative lever. A formulation that raises density while leaving occupancy unchanged moves accessible copy number linearly and accessibility not at all, so effort spent on conjugation chemistry beyond the packing ceiling is wasted, and effort spent on reducing adsorbed protein number is not.
2. The design target should be stated as a coverage with a stated convention. Nominal and realised occupancy differ by up to 28%, and at the reference condition the factor is 1.27; a threshold quoted without its convention can be wrong by a quarter of its value in the direction that flatters the formulation.

3. Ligand class selection carries a coverage budget. A peptide tolerates 0.34 nominal occupancy and is viable at plasma-like loads, a nanobody tolerates 0.22 and is marginal, and an intact antibody tolerates 0.10 and requires a corona-suppressed surface to retain even half its ligands, which makes the widespread practice of displaying full IgGs on EVs a quantitative liability unless the corona is addressed first.

The model has clear limits, and several of them are consequential. Caps are rigid and static, whereas the real corona is dynamic, exchanges between hard and soft components and varies between individuals and disease states [3, 6]; a dynamic corona would partially re-expose ligands and would add a time axis this model does not have. Ligands are placed uniformly, whereas conjugation chemistries produce clustering and heterogeneous single-particle densities, and clustering could either protect a patch of ligands or concentrate them where a single protein occludes several [2]. Ligand orientation is not resolved, and a ligand presented at an unfavourable angle is functionally inaccessible without being sterically covered, which means that the binary metric is an upper bound on functional accessibility rather than a measurement of it. Proteolysis, aggregation and clearance are excluded, and each can reduce effective targeting independently of shielding. The fractional metric, based on point sampling, is a proxy for epitope exposure rather than a calculation of it, and depends on the assumption that any uncovered part of a ligand footprint is chemically available.

The relationship to prior modelling should be stated explicitly. Steric shielding by adsorbed protein has been modelled extensively for synthetic nanoparticles, where corona composition predicts cellular interaction better than particle size, aggregation state or surface charge [7], building on the general account of the corona as the particle's biological identity [8]. That literature establishes the mechanism but parameterises it for hard synthetic surfaces with fixed ligand chemistry, and it generally treats ligand density as a free variable that improves performance. For EVs, corona characterisation [3–5] and corona-aware surface engineering [4] have advanced quickly, but the quantitative link from coverage to retained accessibility has not been expressed as a threshold. The contribution here is that link, in a form compact enough to use during formulation design, together with an explicit statement of its uncertainty and of the convention in which its coverage parameter is expressed. The most consequential difference from the nanoparticle literature is the negative result on density, which that literature's framing does not anticipate.

Three extensions would strengthen the framework. Coupling the accessibility term to an existing EV pharmacokinetic model would convert a static fraction into a time-resolved targeting efficiency and would allow soft-corona exchange to be represented explicitly. Replacing uniform ligand placement with clustered distributions would test whether clustering protects or penalises accessibility, which is currently undetermined and experimentally accessible. Finally, a small set of measurements, quenching of ligand fluorescence by adsorbed protein or single-particle colocalisation of ligand and corona, would anchor the thresholds to a measured accessibility rather than to a geometric prediction, and would also settle whether the binary or the fractional metric corresponds more closely to receptor engagement.

5. Conclusion

Steric shielding of EV-displayed targeting ligands by plasma protein corona is governed by corona coverage and by the ratio of ligand to corona projected area, not by ligand density. Across a hundred-fold density range the binary accessible fraction on a 100 nm vesicle deviated by at most 0.031 and showed no trend, so over-conjugation cannot rescue targeting once a corona has formed; it buys accessible copies at a fixed rate and nothing more.

The governing relation is a closed form, ALF = exp[-f(1 + (S_ligand/S_corona)^(1/2))²], recovered from 280 simulated means with exponents 1.000, 1.000 and 2.002 and a root-mean-square deviation of 0.0083. It is scale-free between 40 and 160 nm to within 8% and implies half-accessibility coverages of 0.34, 0.22, 0.16 and 0.10, in nominal terms, for RVG peptide, nanobody, transferrin and IgG; in realised occupancy these are 0.29, 0.20, 0.15 and 0.10. Sensitivities to a ± 20% ligand footprint and a 10–14 nm² corona footprint stay within − 12% to + 16% and preserve the ordering of the classes.

The design consequence is that corona suppression and ligand size selection carry the burden, that a threshold should always be quoted with its coverage convention, and that the widely used strategy of displaying intact antibodies at high density is the least favourable combination the model admits. The model delimits one mechanism under stated idealisations rather than predicting in-vivo performance, and its thresholds are intended to be refined by measurement of ligand accessibility in plasma.

Declarations

Competing Interests

The author declares no competing interests.

Author Contributions

Zhuofan Shen: conceptualisation; model design; computational implementation; formal analysis; manuscript writing. All computational procedures, parameter values, random seed, replicate counts and evaluation criteria are documented in this manuscript so that the results can be regenerated from the text.

Acknowledgements

The author thanks the International Society for Extracellular Vesicles (ISEV) community for public resources and published datasets that supported this modelling work. No external funding was specifically allocated for this study.

Data Availability

All simulated values underlying the results are reported in Sections 3.2, 3.3, 3.5 and 3.6, and the derived thresholds, fit coefficients and sensitivity intervals in Sections 3.4, 3.6 and 3.7. The model is fully specified by the assumptions, sampling rules and counts in Sections 2.1–2.3 and by Appendix A; the simulation used NumPy and SciPy with the seed 20260614, fifteen replicates per coverage-sweep condition and twenty-five per density-sweep condition. No external repository, dataset or code archive is required for reproduction.

References

1. Wiklander OPB, Nordin JZ, O'Loughlin A et al (2015) Extracellular vesicle in vivo biodistribution is determined by cell source, route of administration and targeting. J Extracell Vesicles 4:26316
2. Pan X, Zhang S, Xia L et al (2025) Unravelling ligand conjugation performance in extracellular vesicles: a quantitative assessment of lipid, protein, and membrane modifications. J Extracell Vesicles 14(10):e70174
3. Tóth EÁ, Turiák L, Visnovitz T et al (2021) Formation of a protein corona on the surface of extracellular vesicles in blood plasma. J Extracell Vesicles 10(11):e12140
4. Musicò A, Zenatelli R, Romano M et al (2023) Assessing the effect of protein corona formation in the process of EV surface engineering. ChemRxiv [preprint]. 10.26434/chemrxiv-2023-rh5s8
5. Ragni E, Taiana M (2024) Mesenchymal stromal cells-extracellular vesicles: protein corona as a camouflage mechanism? Extracell Vesicles Circulating Nucleic Acids 5(4):722–727
6. Liam-Or R, Faruquie SS, Oieni F et al (2024) Cellular uptake and in vivo distribution of mesenchymal-stem-cell-derived extracellular vesicles are protein corona dependent. Nat Nanotechnol 19:846–855
7. Walkey CD, Olsen JB, Song F et al (2014) Protein corona fingerprinting predicts the cellular interaction of gold and silver nanoparticles. ACS Nano 8(3):2439–2455
8. Monopoli MP, Åberg C, Salvati A, Dawson KA (2012) Biomolecular coronas provide the biological identity of nanosized materials. Nat Nanotechnol 7(12):779–786

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryNotesandAppendix.docx