Electronic Supplementary Material (ESI) for Soft Matter.

Electronic Supplementary Information

Complexation of Single Stranded RNA with an Ionizable Lipid: an all-atom Molecular Dynamics Simulation Study

Anastassia N. Rissanou[1,2,*], Andreas Ouranidis[1], Kostas Karatasos[1,*]

1. Department of Chemical Engineering, University of Thessaloniki, P.O. BOX 420, 54124 Thessaloniki, Greece.
2. Department of Mathematics and Applied Mathematics, University of Crete, GR-71409, Heraklion, Crete, Greece.

[*] Author to whom correspondence should be addressed:
risanou@uoc.gr +30 2810393746 fax: +30 2810393701
karatas@eng.auth.gr +30-2310995850 fax: +30-2310996222

Nucleotide sequence for the short RNA fragment

G A U A U U C A G A U G A C C C A G A G C C C G A C A G C

Nucleotide sequence for the long RNA sequence

GAUAUUCAGAUGACCCAGAGCCCGAGCAGCCUGAGCGCGAGCGUGGGCGAUCGCG
UGACCAUUACCUGCCGCGCGAGCCAGGAUGUGAACACCGCGGUGGCGUGGUAUCA
GCAGAAACCGGGCAAAGCGCCGAAACUGCUGAUUUAUAGCGCGAGCUUUCUGUA
UAGCGGCGUGCCGAGCCGCUUUAGCGGCAGCCGCAGCGGCACCGAUUUUACCCUG
ACCAUUAGCAGCCUGCAGCCGGAAGAUUUUGCGACCUAUUAUUGCCAGCAGCAUU
AUACCACCCCGCCGACCUUUGGCCAGGGCACCAAAGUGGAAAUUAAACGCACCGU
GGCGGCGCCGAGCGUGUUUAUUUUUCCGCCGAGCGAUGAACAGCUGAAAAGCGGC
ACCGCGAGCGUGGUGUGCCUGCUGAACAACUUUUAUCCGCGCGAAGCGAAAGUGC
AGUGGAAAGUGGAUAACGCGCUGCAGAGCGGCAACAGCCAGGAAAGCGUGACCG
AACAGGAUAGCAAAGAUAGCACCUAUAGCCUGAGCAGCACCCUGACCCUGAGCAA
AGCGGAUUAUGAAAAACAUAAAGUGUAUGCGUGCGAAGUGACCCAUCAGGGCCU
GAGCAGCCCGGUGACCAAAAGCUUUAACCGCGGCGAAUGC

Identification codes

PDB ID: 5TDN (Home sapiens)

MMDB ID: 145518

A consensus sequence refers to all the possible codons for each amino acid returned from reverse translation engines.

>reverse translation LC Trastuzumab of 5TDN (Tras to a 642 base sequence of consensus codons

gayathcaratgacncarwsnccnwsnwsnytnwsngcnwsngtnggngaymgngtnacnathacntgymgngcnwsncarg
aygtnaayacngcngtngcntggtaycarcaraarccnggnaargcnccnaarytnytnathtaywsngcnwsnttyytntaywsng
gngtnccnwsnmgnttywsnggnwsnmgnwsnggnacngayttyacnytnacnathwsnwsnytncarccngargayttygc
nacntaytaytgycarcarcaytayacnacnccnccnacnttyggncarggnacnaargtngarathaarmgnacngtngcngcncc

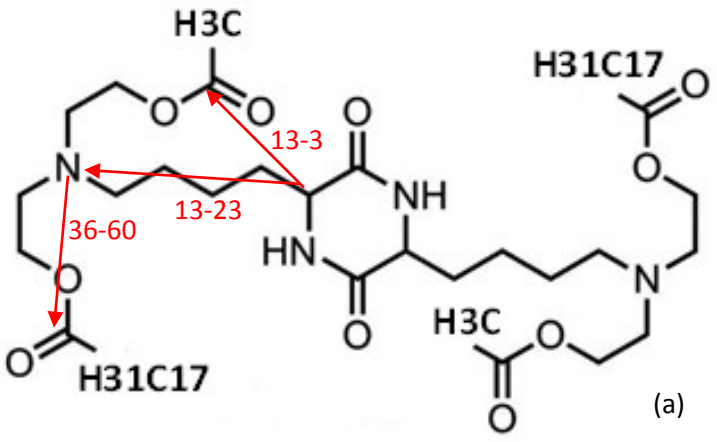
nwsngtnttyathttyccnccnwsngaygarcarytnaarwsnggnacngcnwsngtngtntgyytnytnaayaayttytayccnm
gngargcnaargtncartggaargtngayaaygcnytncarwsnggnaaywsncargarwsngtnacngarcargaywsnaarga
ywsnacntaywsnytnwsnwsnacnytnacnytnwsnaargcngaytaygaraarcayaargtntaygcntgygargtnacncay
carggnytnwsnwsnccngtnacnaarwsnttyaaymgnggngartgy

Results for 214 residue sequence starting "DIQMTQSPSS"
>reverse translation LC Trastuzumab of 5TDN to a 642 base sequence of most likely codons
Gatattcagatgacccagagcccgagcagcctgagcgcgagcgtgggcgatcgcgtgaccattacctgccgcgcgagccaggatgtga
acaccgcggtggcgtggtatcagcagaaaccgggcaaagcgccgaaactgctgatttatagcgcgagctttctgtatagcggcgtgccga
gccgctttagcggcagccgcagcggcaccgattttaccctgaccattagcagcctgcagccggaagattttgcgacctattattgccagcag
cattataccaccccgccgacctttggccagggcaccaaagtggaaattaaacgcaccgtggcggcgccgagcgtgtttatttttccgccga
gcgatgaacagctgaaaagcggcaccgcgagcgtggtgtgcctgctgaacaacttttatccgcgcgaagcgaaagtgcagtggaaagtg
gataacgcgctgcagagcggcaacagccaggaaagcgtgaccgaacaggatagcaaagatagcacctatagcctgagcagcaccctga
ccctgagcaaagcggattatgaaaaacataaagtgtatgcgtgcgaagtgacccatcagggcctgagcagcccggtgaccaaaagcttta
accgcggcgaatgc

>transcription LC Trastuzumab of 5TDN to a 642 base sequence of most likely codons derived from consensus codons Homo Sapiens
GAUAUUCAGAUGACCCAGAGCCCGAGCAGCCUGAGCGCGAGCGUGGGCGAUCGCG
UGACCAUUACCUGCCGCGCGAGCCAGGAUGUGAACACCGCGGUGGCGUGGUAUCA
GCAGAAACCGGGCAAAGCGCCGAAACUGCUGAUUUAUAGCGCGAGCUUUCUGUA
UAGCGGCGUGCCGAGCCGCUUUAGCGGCAGCCGCAGCGGCACCGAUUUUACCCUG
ACCAUUAGCAGCCUGCAGCCGGAAGAUUUUGCGACCUAUUAUUGCCAGCAGCAUU
AUACCACCCCGCCGACCUUUGGCCAGGGCACCAAAGUGGAAAUUAAACGCACCGU
GGCGGCGCCGAGCGUGUUUAUUUUUCCGCCGAGCGAUGAACAGCUGAAAAGCGGC
ACCGCGAGCGUGGUGUGCCUGCUGAACAACUUUUAUCCGCGCGAAGCGAAAGUGC
AGUGGAAAGUGGAUAACGCGCUGCAGAGCGGCAACAGCCAGGAAAGCGUGACCG
AACAGGAUAGCAAAGAUAGCACCUAUAGCCUGAGCAGCACCCUGACCCUGAGCAA
AGCGGAUUAUGAAAAACAUAAAGUGUAUGCGUGCGAAGUGACCCAUCAGGGCCU
GAGCAGCCCGGUGACCAAAAGCUUUAACCGCGGCGAAUGC

Autocorrelation Functions

The orientational dynamics of the DML molecules belonging in clusters can be studied through the autocorrelation function of the second Legendre polynomial, $P_2(\cos\theta(t)) = \frac{3}{2}\left\langle \cos^2\theta(t) \right\rangle - \frac{1}{2}$, for vectors defined along the molecule (as described in Figure SI-1a in the Supporting Material). The angle θ refers to the angle between a defined direction at time t=0 and the same direction at time t. As time lapses the orientation of a vector along the selected direction changes, since parts of the molecule (or the entire molecule under examination) changes its orientation. For this purpose 3 kinds of vectors (representing 3 different directions) have been selected; one originating from the ionizable nitrogen atom to the end carbon atom of the H31C17 arm, referred to as "36-60", a second one starting from a carbon atom of the phenyl ring to the end carbon of the H3C end group, referred to as "13-3" and a third one starting from a carbon atom of the phenyl ring and pointing towards the ionizable nitrogen atom, referred to as "13-23". Since there are two identical vectors of each kind due to the symmetry of the DML lipid, autocorrelation functions have been calculated as averages over them. The corresponding results are presented in Figure SI-1b, where it is demonstrated that for all three vectors along the examined directions, decorrelation is achieved at a timescale close to 20-30 ns. Therefore, during the production runs with length covering about 100ns, the sampling of the orientational configurational space of the complexation agent is sufficient.



(a)

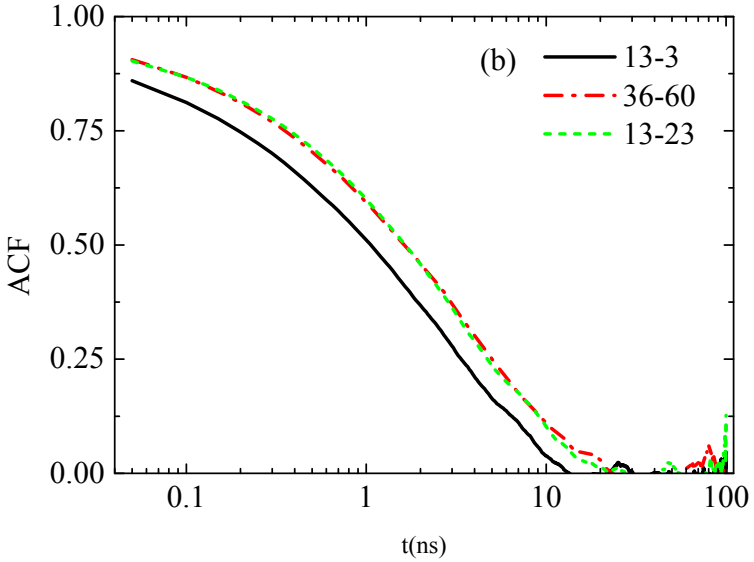
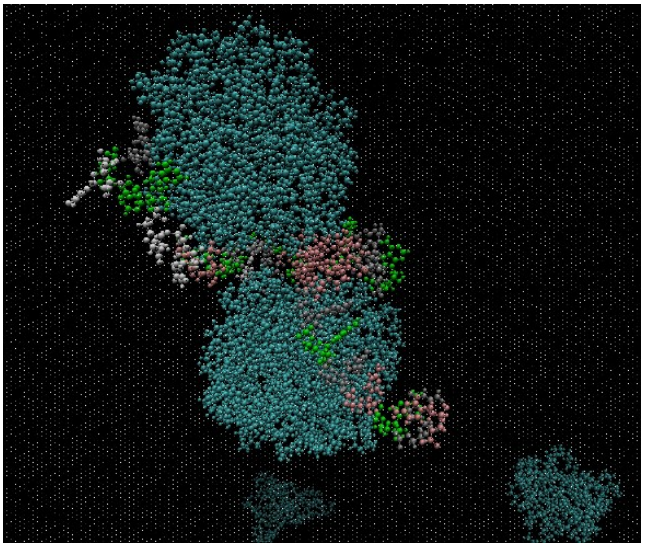


Figure SI-1: (a) Chemical structure of DML and vector definition. (b) Autocorrelation function of the second Legendre polynomial for vectors "13-3", "36-60" and "13-23" along the DML molecule.

DMLRW system at Ionic Strength of 0.15M

(a)



(b)

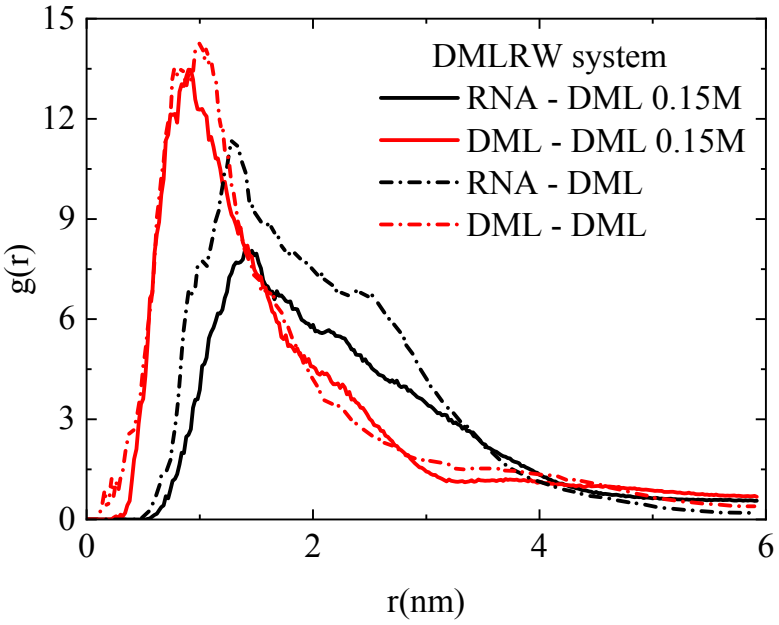


Figure SI-2: (a) DML clusters complexed with the oligonucleotide RNA in a DMLRW-analogous system at an ionic strength of 0.15M. Water molecules are presented as ghost molecules for clarity. DML molecules appear in dark cyan. Atoms of different colors belong to RNA. (b) Pair radial distribution function between the centers of mass of RNA – DML and DML – DML for DMLRW system and the analogous system at 0.15M ionic strength.

Histograms of the cluster size

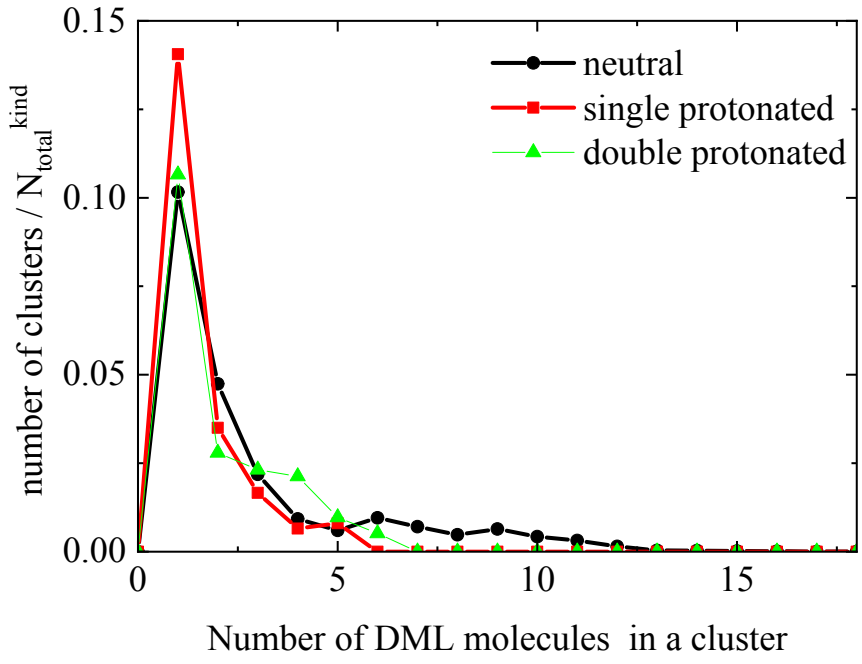


Figure SI-3: Histogram of the number of clusters, in terms of the number of DMLs they contain, for clusters formed by the 3 different kinds of DMLs.

In Figure SI-3 cluster histograms for each of the different kinds of DMLs are shown, in order to check for any particular differentiation. Since a different number of DMLs of each kind have been used in the simulation, the cluster population was normalized with this number. Qualitatively similar distributions are observed for the three kinds of clusters. However, clusters of protonated DMLs are comprised by a smaller number of molecules compared to those formed by neutral DMLs, as it is can be inferred from the extended tail of the black curve.

Van-Hove correlation functions

Although the DML molecules participating in a cluster remain practically confined within the cluster's volume once the clusters are formed, internal cluster restructuring may take place within a finite timescale. To get an estimation of this timescale, one can probe the collective motion of DML particles and estimate the time it takes for the position of the center of mass of

an DML molecule to become uncorrelated with that of a different one. This analysis can be performed by monitoring the distinct Van-Hove correlation function $G_d(r,t)$[1]:

$$G_d(r,t) = \frac{1}{N}\left\langle \sum_i \sum_{j \neq i} \delta\left[r - \left| \boldsymbol{r}_i(t) - \boldsymbol{r}_j(0) \right| \right] \right\rangle$$

where N is the number of DML particles that are examined, δ represents the Dirac's function and $\boldsymbol{r}_i(t)$ is the position vector of the center of mass of the i_{th} DML particle at time t. The distinct Van Hove function essentially probes density fluctuations due to the collective motion of the examined molecules at a timescale t and at a length scale defined by a distance r between them. At t=0, $G_d(r,t)$ reflects the local arrangement of molecules as described by the radial distribution function $g(r)$, $G_d(r,0)=\rho g(r)$ (ρ is the number density of the probed particles, here represented by the centers of mass of the DMLs). As time lapses, the position of each DML molecule becomes uncorrelated to the initial position of a different such particle, so that $G_d(r,t)$ converges to the average density of the examined molecules. Figure SI-4a illustrates the distinct van Hove correlation functions for different timescales, for the population of the neutral DMLs.

It is shown that as time passes the intensity of the main peak corresponding to the first neighbors at t=0, reduces. This is indicative of a gradual loss of memory of the initial local arrangement assumed by the DMLs at t=0. At long enough timescales, the final structure becomes uncorrelated to the starting configuration and the main peak vanishes. The longer it takes for this peak to lose its amplitude, the longer the lifetime of the local structure. A visual inspection of Figure SI-4a indicates that it takes about 50ns before an internal cluster restructuring takes place. A similar picture describes the behavior of the charged DMLs as well (Figure SI-4b and Figure SI-4c).

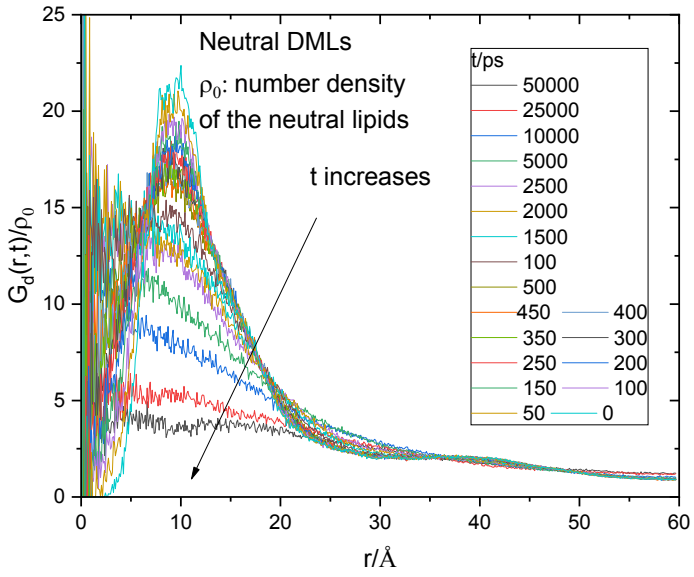


Figure SI-4a: Distinct van Hove function of the neutral DMLs. Here, all neutral DMLs were taken into account since practically all participate in the formed clusters.

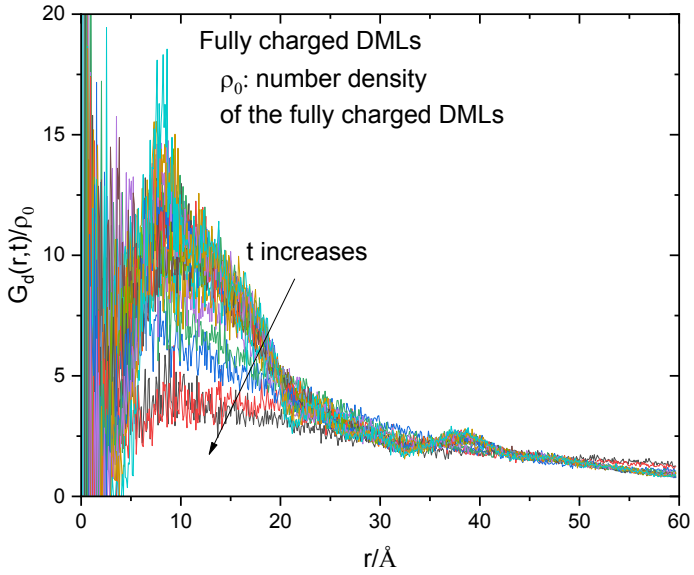


Figure SI-4b: Distinct van Hove function of the fully charged DMLs. Here, all the doubly charged DMLs were taken into account since practically all participate in the formed clusters. The color code corresponding to different timescales is the same as in figure SI-4a.

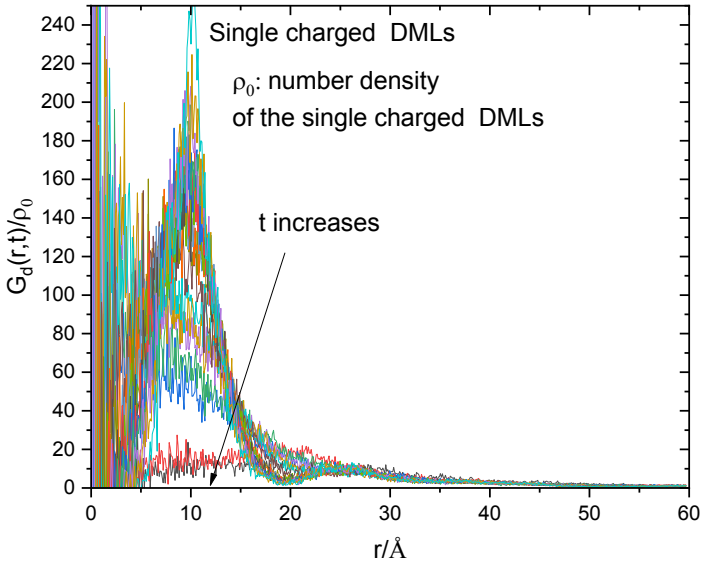


Figure SI-4c: Distinct van Hove function of the single charged DMLs. Here, all the singly charged DMLs were taken into account since practically all participate in the formed clusters. The color code corresponding to different timescales is the same as in figure SI-4a.

Error Bars

Based on the above discussion, we have considered that parts of the single, but long, trajectory are statistically equivalent to independent different runs of shorter length, as far as it concerns the conformational properties of the complexation agent. On this basis, taking different parts of the entire trajectory, we performed calculations based on block averages and estimated error margins. Results for the density and the charge distributions are presented in Figure SI-5.

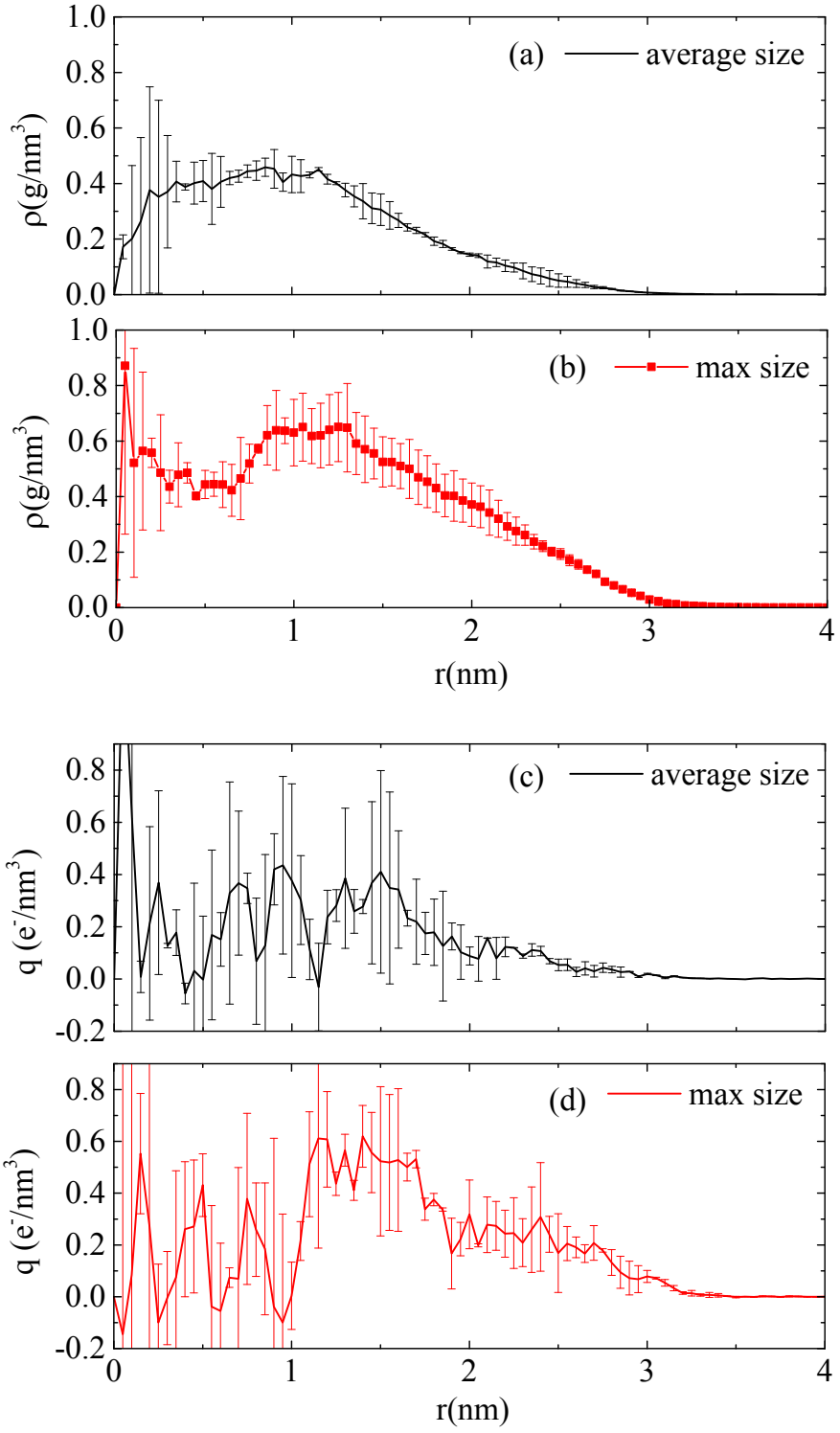


Figure SI-5: (a,b) Density profile and **(c,d)** charge profile, as a function of the distance from the center of the cluster, for a highly and an average populated cluster respectively.

1. J.-P. M. Hansen, I. R., Theory of simple liquids, Elsevier: Amsterdam **3rd ed.** (2006).