Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information

Adsorbate Interactions on GaN(0001) Surface and Their Effect on Diffusion Barriers and Growth Morphology

Manjusha Chugh and Madhav Ranganathan*

Department of Chemistry, Indian Institute of Technology Kanpur, Kanpur. India.

E-mail: madhavr@iitk.ac.in

- To whom correspondence should be addressed.

The DFT results presented here have been obtained within the Generalized Gradient Approximation (GGA) parameterized by PBE exchange-correlation functional using PWscf code of Quantum Espresso package (version 5.0.2). For the optimization calculations, a kinetic energy cutoff of 70 Ry is used and the force convergence threshold was set up at 0.025 eV/Å. Sample input and output files used are provided at the end of this document.
We have performed two types of calculations, "full-relax" and "fixed-slab" calculations. In a full-relax calculation, we start with some initial coordinates and geometry optimization is performed to relax the coordinates. This effectively finds a local minimum in energy around the initial coordinates. In a fixed-slab calculation, first the geometry optimization of a clean slab is done to obtain the optimum atomic coordinates of the substrate atoms. Then these substrate coordinates are kept fixed during the geometry optimization and only the adatom is allowed to relax completely.

S1.1. Convergence tests for slab thickness

We chose a 4 GaN bilayers (BLs) (or 8 layers) slab for our surface calculations with pseudohydrogen termination at the bottom end of the slab. We fixed the bottom 2 GaN BLs (4 layers) and the pseudohydrogen layer to mimic bulk like behavior and allowed the top 2 GaN BLs (4 layers) to relax during the geometry optimization calculations. To confirm that the slab thickness is adequate, we performed the following tests.

We calculated adsorption energies of Ga and N adatoms for slabs of varying thicknesses. Adsorption energy of a Ga adatom on a (5x5) unit cell is -5.38 and -5.47 eV for a 4 BLs and a 6 BLs slab having top 2 and top 3 BLs relaxed respectively. The adsorption energy of a N adatom on a (5x5) unit cell is -7.08 and -7.20 eV for 4 BLs and 6 BLs slab respectively. Thus, we see that adsorption energies do not change much after 4 BLs.
The adsorption energies of a Ga adatom on a (2x2) unit cell are -3.82, -3.83, -3.83 and -3.83 eV for 4, 5, 6 and 7 BLs slabs respectively.

We have calculated the interlayer distances (d_{ij}, where i and j represent successive slab layers) of the slab in the z-direction for adatom pair configurations for the (5x5) unit cells. We found that the interlayer spacings rapidly approach the bulk value as we approach the second bilayer from top. For example, the average interlayer spacings of the top three layers of the slab for a (0,2) Ga-Ga adatom configuration (see Fig. 1 of the manuscript) are found to be 0.63 Å, 1.99 Å and 0.65 Å. The corresponding values of bulk GaN are 0.64, 1.97 and 0.64 Å respectively. Further, the displacement maps shown in Fig. 3 of the manuscript confirm that the atomic displacements become small for atoms in the 2nd BL.

In addition, we have also performed calculations on a 4 BLs slab keeping only the bottom bilayer fixed and relaxing top 3 BLs. The adsorption energy for a Ga adatom in this case is -5.39 eV.
Since we are primarily interested in trends in interaction energies, we feel that the 4 BLs slab with bottom 2 BLs fixed is sufficient for our purpose.

Relaxing the 3rd BL also helps us in estimating the dipole effect of our model. To estimate the same, we calculated the dipole moments of 4 BLs slab with 2 and 3 BLs relaxed respectively. We found that the dipole moment for the former case is -23.65 D and for the latter case, it is -22.95 D, where D stands for

debye. These dipole moments differ by only 0.0035 D/atom.

All the above tests confirm that the 4 GaN BLs slab is sufficient for the calculations performed in this article.

S1.2. Convergence tests for vacuum gap

For the surface calculations, we have chosen a vacuum gap of 13 Å. We have checked whether this gap is sufficient by performing the following tests.
We have calculated adsorption energies of a Ga adatom on a (2x2) unit cell slab having 6 BLs with different vacuum gaps . The adsorption energy of a Ga adatom on a (2x2) unit cell with 13 Å gap is -3.83 eV and with 16 Å gap is -3.83 eV respectively. The adsorption energy of a N adatom on a (2x2) unit cell having a vacuum gap of 13 Å is -5.67 eV and on that having a vacuum gap of 16 Å is -5.66 eV respectively.

We have calculated interlayer distances (d_{ij}) of the slab along the z-direction for the Ga adatom configuration for the (2x2) unit cell having 13 Å and 16 Å gaps. The average interlayer distances d_{12} and d_{23} for the 13 Å gap slab are 0.672, 1.976 and for the 16 Å gap slab are 0.673, 1.975 respectively. All these distances are in Å. The average distance of the Ga adatom from the topmost substrate layer is 1.806 and 1.806 Å for the 13 Å slab and for the 16 Å slab respectively.

From these tests, we can say that the 13 Å vacuum gap is sufficient for doing the surface calculations for GaN(0001). Also note that the bottom side of the slab is passivated by pseudohydrogens, which prevents the two sides of the slab from interacting with each other.

S2. Convergence tests for k-points for different supercells

We have used (2x2) and (3x3) supercells for diffusion barrier calculations and (5x5) supercell for adatom interactions calculations in our work. We have used (NxNx1) k-points for these supercells, which are (4x4x1), (3x3x1) and (1x1x1) k-points for (2x2), (3x3) and (5x5) supercells respectively. We have taken single k-point along the z-direction for all the supercells, as is usually done for the slab geometries having sufficient vacuum gap. Since a particular unit cell has same lateral dimensions, the k-points (represented by a number N) along the lateral dimensions will be same for a particular unit cell. The k-points along the lateral dimensions for different unit cells should be different and should have the same level of accuracy for each unit cell considered.
We have calculated the self consistent field (scf) energy for different k-points along the lateral dimensions for each unit cell and the plots are shown in Fig. S1. For consistency, we have also included (4x4) unit cell and tested the k-point convergence for this also.

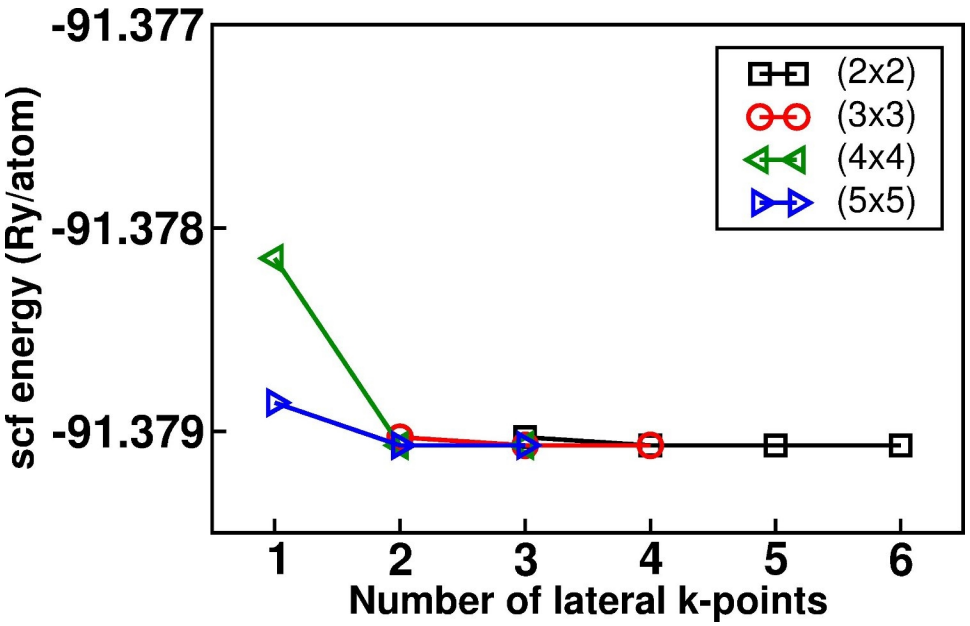

Fig. S1: Self consistent field (scf) energy (Ry/atom) for different supercells as a function of number of k-points along one of the lateral dimensions.

Based on the above results for the convergence tests for k-points for different unit cells, we can say that we should use (4x4x1) and (3x3x1) k-points for (2x2) and (3x3) unit cells respectively. (2x2x1) k-points are good for a (4x4) GaN unit cell. For a (5x5) unit cell, it seems that (2x2x1) k-points are good. But if we want to have same level of accuracy for different unit cells, then we should use (1x1x1) k-points for the (5x5) unit cell. These sets of k-points are consistent in the obtained energetics for these unit cells. So, the appropriate set of k-points for the above unit cells are:

unit cell	k-points
(2x2)	(4x4x1)
(3x3)	(3x3x1)
(4x4)	(2x2x1)
(5x5)	(1x1x1)

S3. Potential distribution along z-direction

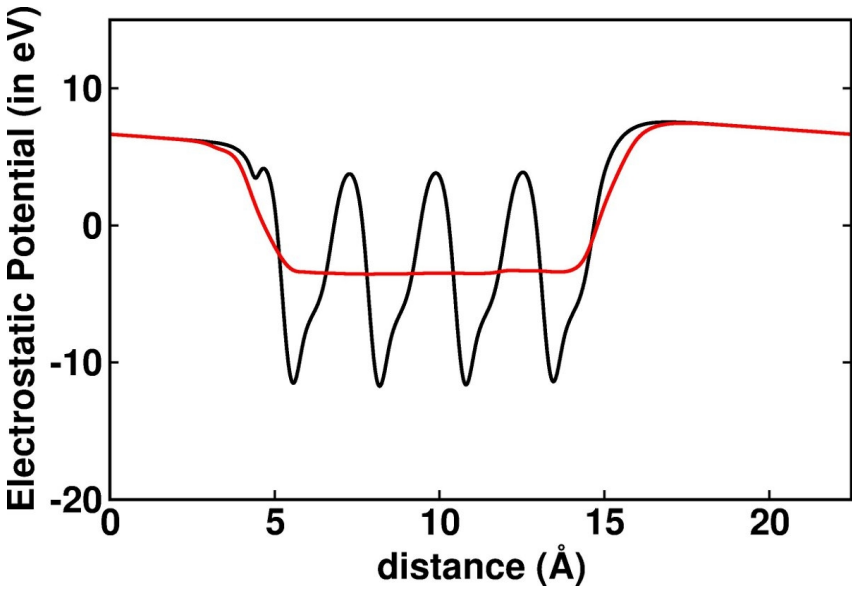

Fig. S2(a): Planar average (black line) and macroscopic average (red line) electrostatic potential along z-direction for a 4 BLs thick clean (5x5) GaN slab.

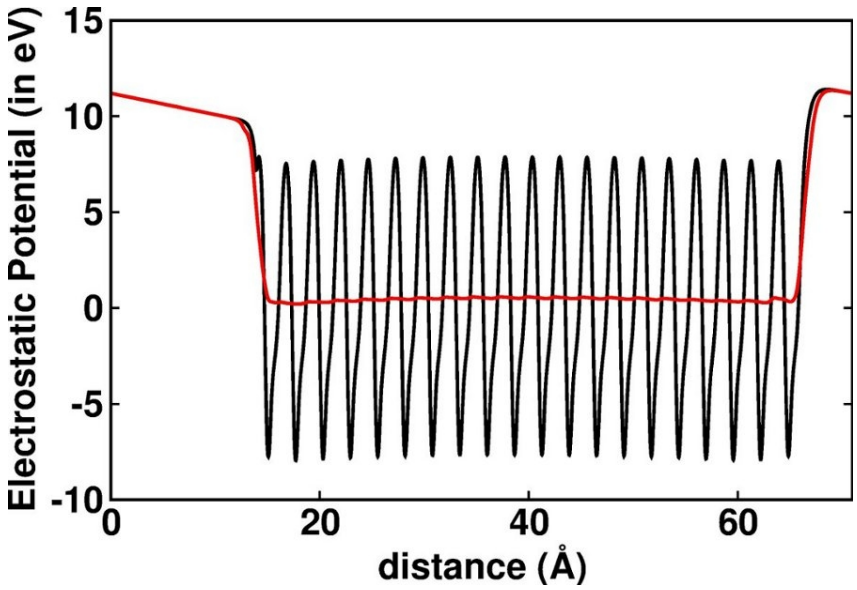

Fig. S2(b): Planar average (black line) and macroscopic average (red line) electrostatic potential along z-direction for a 20 BLs thick clean (1x1) GaN slab.

S4. Band structure calculations

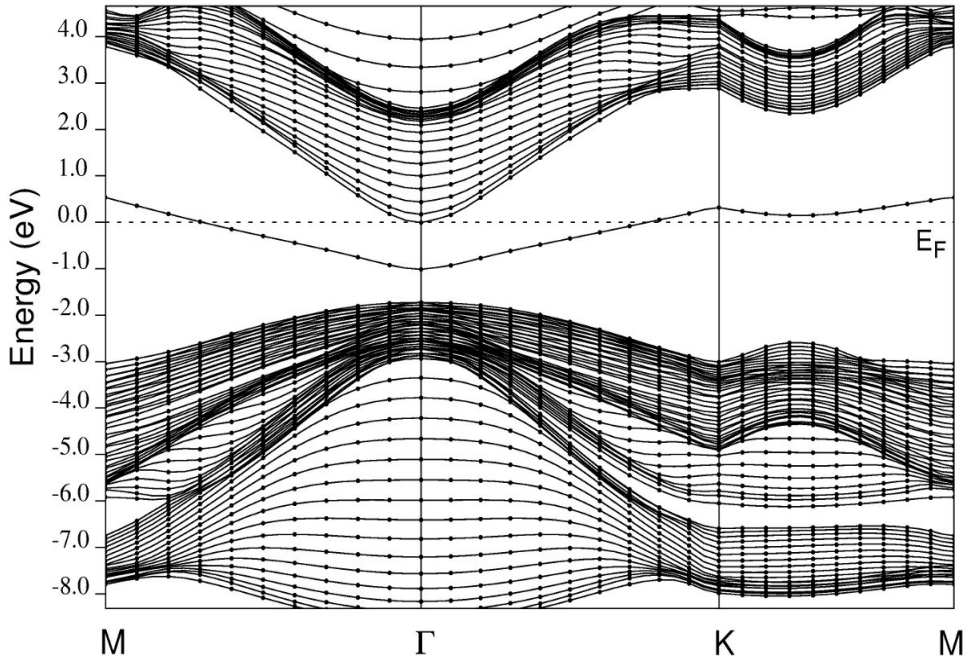

Fig. S3: Band diagram (dispersion relation) for a 20 BLs clean (1x1) slab terminated by pseudohydrogens of charge 0.75e at the optimized distance of 1.03 Å from the bottom N-layer. The energy is shown on y-axis and the wave vector is shown on x-axis.

S5. Adsorption energy of Ga for different pseudohydrogen charges

We calculated the adsorption energy of a Ga adatom on a (5x5) unit cell for pseudohydrogens of charges 0.50e, 0.75e and 1.0e. The adsorption energy values for these respective cases are -5.91 eV, -5.38 eV and -4.84 eV respectively. We have optimized the pseudohydrogens distance from the bottom N-layer in all the three cases above.

S6.1 FCS and DCS configurations and their energetics

A clean slab is obtained from the bulk GaN crystal by cutting it at Ga-termination perpendicular to the [0001] direction. This slab remains flat after geometry optimization of top few bilayers. We refer to this flat slab as a Flat Clean Slab (FCS). Fig. S4(a) refers to a FCS configuration. After adsorption of adatom/adatoms on this FCS and after geometry optimization, another configuration is obtained, as shown in in Fig. S4(c). We refer to this as adatoms/substrate system.

A clean slab can also be obtained by removing the adsorbed species from the adsorbate/substrate configuration, as shown in Fig. S4(d). Since the atoms in this configuration are displaced (distorted) as compared to those in FCS, we refer to this as a Distorted Clean Slab (DCS) configuration.
The energy of this DCS can be obtained in two ways. One is by just doing the self consistent field (scf) calculation of the DCS configuration. The clean slab energy obtained in this way is called DCS-scf energy. Other way to obtain the clean slab energy is by performing a geometry optimization of the DCS configuration. The configuration obtained after this is called a DCS-rlx configuration (shown in Fig. S4(e)) and the energy of this clean slab is called DCS-rlx energy.

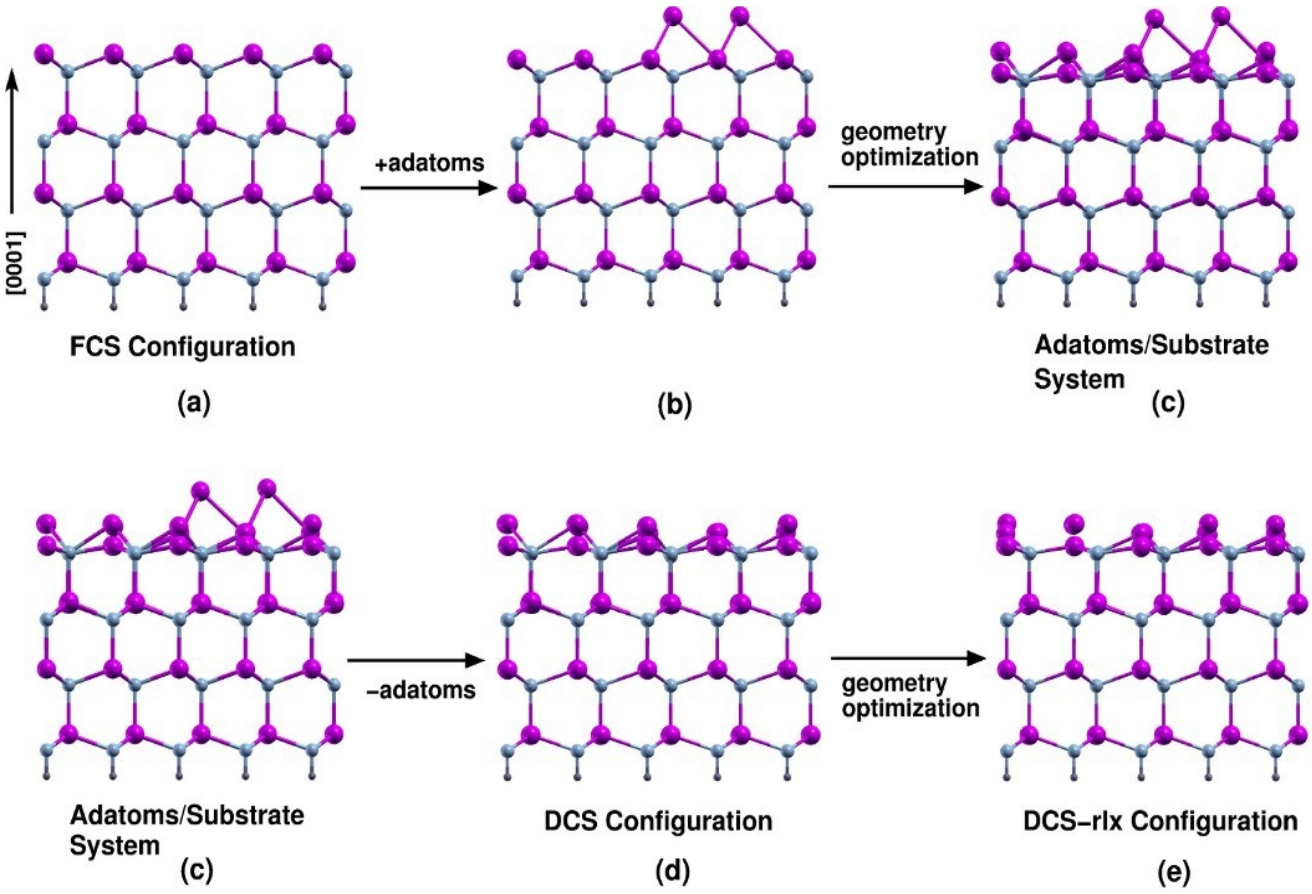

Fig. S4: GaN(0001) surface configurations used in this work. FCS represents a flat clean slab. DCS represents a distorted clean slab (DCS). Large pink colored spheres denote Ga atoms, small light blue spheres denote N atoms and smallest gray spheres denote H atoms of the pseudohydrogen layer respectively.

In our manuscript, we have used the energy of the FCS and DCS-rlx configurations as a reference for the clean slab to calculate the interaction energy between the adatoms. These are represented in equations 1 and 5 of the manuscript, which are also written below.

$$E_{\text{int}}^{AB} = E_{slab}^{AB/GaN} - E_{slab}^{A/GaN} - E_{slab}^{B/GaN} + E_{slab}^{GaN} \qquad (1)$$

$$E_{\text{int}}^{AB'} = E_{LEslab}^{AB/GaN} - E_{LEslab}^{A/GaN} - E_{LEslab}^{B/GaN} + E_{DCS-rlx}^{GaN} \qquad (2)$$

The meanings of each symbol above have been explained in the manuscript.

It is to be noted that for (5x5) clean GaN(0001) surface, there are many DCS-rlx configurations possible. These all are lower in energy than the FCS. Out of all the DCS-rlx configurations for the (5x5) surface unit cell, the lowest energy configuration corresponds to the one in which we start with two Ga adatoms at (0,3) sites, remove the adatoms and optimize the clean slab. For a reference, we give the total energies of the FCS and the DCS-rlx configuration corresponding to Ga-Ga (0,3) pair. The energy of the (5x5) supercell FCS configuration is -88.13487 Ry/Å^2 whereas the energy of the (5x5) supercell Ga-Ga (0,3) DCS-rlx configuration is -88.13576 Ry/Å^2. Thus, the energy difference between the FCS and most stable DCS-rlx configuration is almost 2.6 eV for a (5x5) unit cell. This energy difference is reflected in the calculated interaction energies for the adatom pairs using Eq. 1 and 2 above (see Figs. 4 and 5 of the manuscript). Similarly, for other adatom pairs configurations, the DCS-rlx has lower energy than the FCS. These energies are tabulated below.

DCS-rlx Configurations	**Energy with ref. to FCS (eV/surface Ga atom)**
Ga-Ga (0,1)	-0.07
Ga-Ga (0,2)	-0.10
Ga-Ga (0,3)	-0.11
Ga-Ga (0,4)	-0.11
N-N (0',1')	-0.09
N-N (0',2')	-0.06
N-N (0',3')	-0.10
N-N (0',4')	-0.10
Ga-N (0,0')	-0.10
Ga-N (0,1')	-0.10
Ga-N (0,2')	-0.08
Ga-N (0,3')	-0.07
Ga	-0.07
N	-0.07

S6.2 Interaction energies using DCS-rlx

We took the most stable DCS-rlx configuration of the (5x5) supercell, i.e. the one obtained from the Ga-Ga (0,3) configuration. We placed the adatom/adatom pairs again at their binding sites at this DCS-rlx configuration and performed the geometry optimizations. The total energies of these configurations are represented as $E_{DCSslab}^{AB/GaN}$ for adatom pair, $E_{DCSslab}^{A/GaN}$ and $E_{DCSslab}^{B/GaN}$ for single adatom configurations, where A and B can be either Ga or N.

The interaction energy between the adatoms for these configurations can be calculated as:

$$E_{\text{int}}^{AB,DCS-rlx} = E_{DCSslab}^{AB/GaN} - E_{DCSslab}^{A/GaN} - E_{DCSslab}^{B/GaN} + E_{DCS-rlx}^{GaN} \qquad (3)$$

where $E_{DCS-rlx}^{GaN}$ is the energy of the most stable DCS-rlx configuration of the clean (5x5) sufrace unit cell. The interaction energies obtained from equation 3 are plotted with respect to the intersite distance between the adatoms and are shown in Fig. S5.

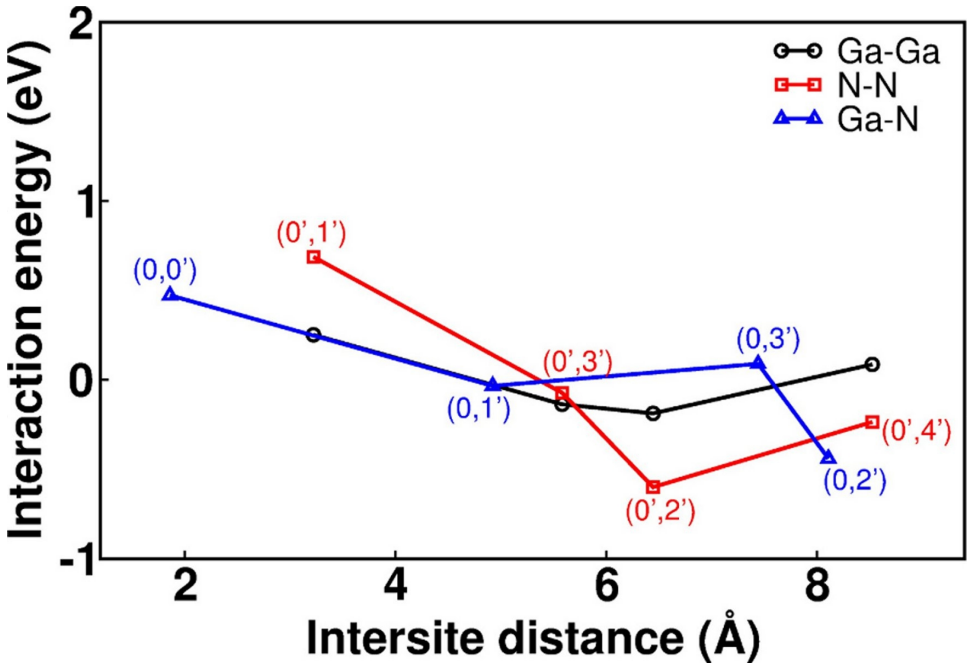

Fig. S5: Interaction energies between Ga-Ga, N-N and Ga-N adatom pairs, calculated using the DCS-rlx configurations. The N-N and Ga-N pairs are labeled.

Note that these interaction energies graphs are different from that shown in Fig. 5 of the manuscript, which is calculated according to equation 2 above or equation 5 of the manuscript and where the lowest energies of the adatom/adatoms pairs configurations were used.

S7. Adatom distances from the substrate

	Pairs	d_{Ga1-S}	d_{Ga2-S}	d_{N1-S}	d_{N2-S}	d_{Ga-S}	d_{N-S}
Ga-Ga	(0,1)	2.16	2.16				
	(0,2)	1.84	1.84				
	(0,3)	1.86	1.97				
	(0,4)	1.84	1.83				
N-N	(0',1')			1.37	1.37		
	(0',2')			1.27	1.27		
	(0',3')			1.27	1.28		
	(0',4')			1.27	1.26		
Ga-N	(0,0')					3.20	1.28
	(0,1')					1.84	1.26
	(0,2')					1.97	1.27
	(0,3')					1.84	1.22

Table S1: Distances of adatoms from the average height of the substrate. Each adatom for a particular pair configuration for Ga-Ga and N-N pairs is numbered as 1 and 2. 'S' represents the substrate. The distance of Ga adatom from the substrate is quite large in case of (0,0') Ga-N pair configuration. This is because in this case, the Ga adatom goes nearly on-top of the neighboring N adatom after geometry optimization. Hence the final distance of the Ga adatom from the substrate is more. The labelling of different pairs are according to figure 1 of the manuscript. For convenience, we have shown the figure 1 of the manuscript below.

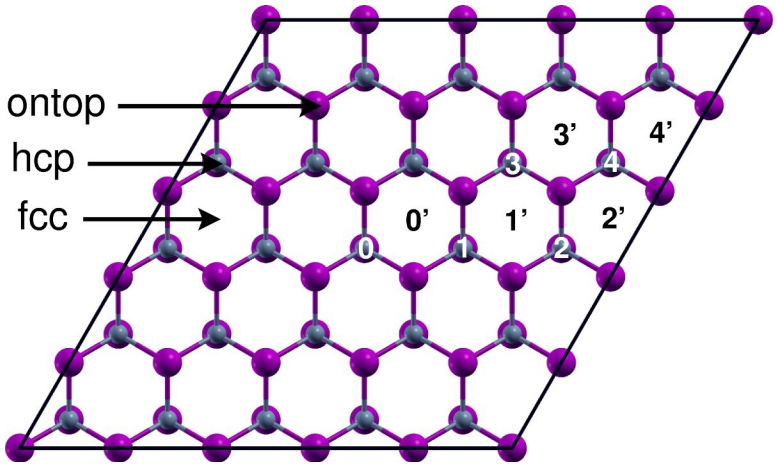

Fig. S6 : Top view of GaN(0001) surface showing a (5x5) supercell. Adatoms sites are labelled as

numbers. Ga-Ga pairs are : (0,1), (0,2), (0,3) and (0,4); N-N pairs are : (0',1'), (0',2'), (0',3') and (0',4'); Ga-N pairs are : (0,1'), (0,2'), (0,3') and (0,4'). This figure is same as figure 1 of the manuscript.

S8. Charge calculations

We have calculated Bader charges on Ga and N atoms in the fixed-slab and full-relax calculations[1,2]. Here, we show the Bader charges for two adatoms and top two layers of the (0,2) Ga-Ga pair configuration (see Fig. 1 of the manuscript).

Atom No.	Bader Charge full-relax	Bader Charge fixed-slab
176	6.5444	6.5914
177	6.5735	6.5944
178	6.5714	6.5916
179	6.5468	6.5909
180	6.5694	6.5916
181	6.5766	6.5886
182	6.5707	6.5906
183	6.5753	6.5859
184	6.5238	6.5872
185	6.5228	6.5850
186	6.5278	6.5877
187	6.5288	6.5864
188	6.5644	6.6215
189	6.5967	6.5859
190	6.5627	6.6227
191	6.5294	6.5890
192	6.5898	6.5888
193	6.5352	6.5874
194	6.5344	6.5889
195	6.5867	6.5904
196	6.5429	6.5946
197	6.5516	6.5923
198	6.5667	6.5944
199	6.5522	6.5923
200	6.5390	6.5947
201	11.4617	11.8058
202	12.3917	11.8090
203	11.4564	11.8014
204	12.3918	11.8094
205	11.4611	11.8053
206	11.4683	11.7531
207	11.9843	11.8983
208	11.4670	11.7485
209	12.3969	11.8066
210	11.4549	11.7898
211	11.4547	11.7923
212	11.4609	11.7537
213	12.3649	11.7732

214	12.3684	11.7757
215	11.9831	11.9002
216	11.9215	11.8895
217	11.9223	11.8901
218	11.4680	11.7520
219	12.0109	11.9038
220	11.4628	11.6930
221	12.0114	11.9013
222	11.4609	11.7923
223	11.9819	11.7940
224	11.9785	11.7900
225	11.4688	11.7944
226	12.9683	13.0268
227	12.9678	13.0269

Atom numbers 176-200 represent atoms in the substrate N layer, 201-225 represent atoms in the topmost substrate Ga layer and 226-227 represent two Ga adatoms respectively.
We remind the reader that the 3d-electrons of Ga are treated as valence electrons, so the Bader charge on an isolated Ga atom is 13 and that on N is 5.

S9.1. Dipole moment values for various adatom pairs configurations

We have calculated dipole moments of various adatom pairs configurations ($\mu_{AA/GaN}$) (see Fig. 1 of the manuscript) and of the corresponding clean slab configurations (DCS) (μ_{DCS}) according to equation 3 of the manuscript. A can be either Ga or N. The DCS configuration is described in section S6.1 and also in the manuscript. Here, we show the values of various dipole moments.

	Configurations	$\mu_{AA/GaN}$	μ_{DCS}
Ga-Ga	(0,1)	-25.50	-24.89
	(0,2)	-22.65	-25.11
	(0,3)	-23.35	-26.01
	(0,4)	-23.30	-25.59
N-N	(0',1')	-24.27	-21.61
	(0',2')	-25.85	-23.72
	(0',3')	-25.47	-23.77
	(0',4')	-25.46	-23.56
Ga-N	(0,0')	-24.66	-24.81
	(0,1')	-24.77	-24.40
	(0,2')	-23.37	-23.92
	(0,3')	-23.65	-23.39

Table S2: Dipole moments for various adatom pairs configurations for the full-relax calculations, calculated according to equation 3 of the manuscript.

The dipole moment values are in Debye. Refer to Fig. 1 of the manuscript or Fig. S6 for various configurations' labels.

S9.2. Dipole moment calculations by work function and charge transfer methods

Further, in the manuscript, we have shown dipole moments calculated according to Eq. 3. Here, we show other methods to calculate the dipole moment.

We first calculated the amount of charge transfer to/from the substrate from/to the adsorbed adatoms using the Bader charge analysis method. Then by taking the product of that charge transfer and the distance between the adsorbed species and the substrate, we calculated the dipole moment. The change in dipole moment for a (0',2') pair of N adatoms on a (5x5) unit cell for full-relax case is -0.67 Debye, while this change in dipole moment calculated by Eq. 3 of manuscript is -2.13 Debye (see Table S2). Thus, the dipole moment calculated using the charge transfer method is also small, indicating small dipole-dipole interaction between the adsorbed species.

We also calculated the dipole moment of a (2x2) N adatom configuration. The dipole moment change according to Eq. 3 of manuscript is -2.10 D, while the same calculated using the charge transfer method, it is -2.06 D.

We also calculated the change in dipole moment upon N adatom adsorption using the work function change method. In this method, first the work function of a system is calculated both with and without adatom and their difference is taken. This work function change (ΔW) is then used to calculate the change in dipole moment according to the following Helmholtz equation[3] :

$$\Delta \mu = \frac{A \Delta W}{12\pi\theta}$$

where A is the area of (1x1) unit cell in $Å^2$, θ is the coverage.

The calculated change in dipole moment according to the work function method for the (2x2) N-adatom is -1.31 D.

From all these methods, we can see that the change in dipole moment upon adatom adsorption is very small and thus the calculated dipole-dipole interaction energy is also very small.

S10. Charge density difference map

We calculated the change in charge density upon adatom adsorption according to Eq. 2 of the manuscript. The charge density difference maps show the charge redistribution taking place when an adatom is adsorbed. Here, we show the charge density difference map (Fig. S7) for a N adatom adsorbed at the fcc site on a (5x5) unit cell. The plane of the plot passes through the plane formed by 3 surface Ga atoms bonded to the N adatom. In this plot, we see the positive (blue/pink) and negative

(red) charge regions which correspond to charge accumulation and depletion respectively.
We see that the typical charge density oscillations visible in some adatom/metal (and adatom/graphene) [4] systems are absent in this plot. The charge density oscillations (Friedel Oscillations) present on a metal (and graphene) surface are visible as alternating positive and negative charge regions with decaying amplitude as we go away from the adatom location. But here, we see that the major charge redistribution is present around the adatom location and there are no oscillations visible farther apart from the adatom location (at least with-in this range of charge density). Thus, we can say that the substrate-mediated electronic interactions (which give rise to charge density oscillations) are weak at GaN(0001) surface.

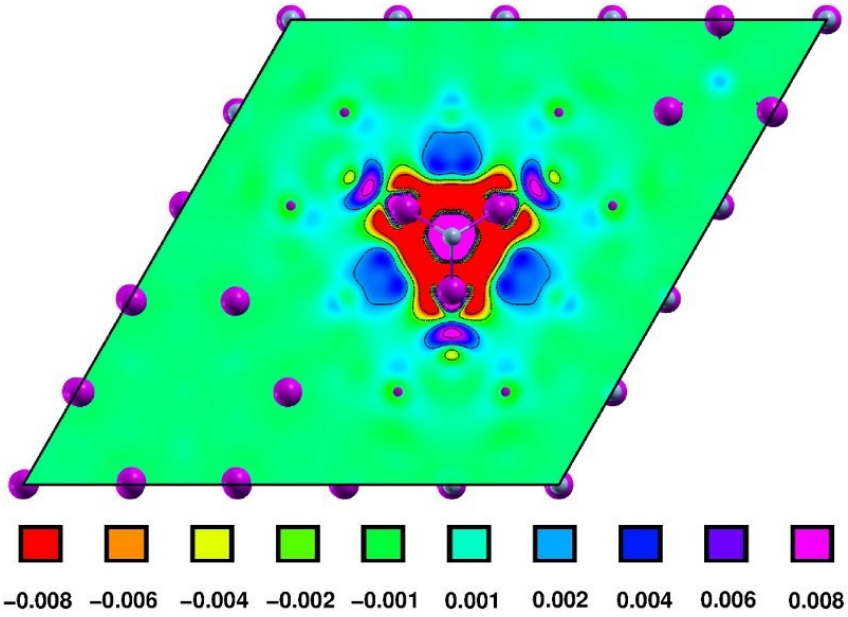

Fig. S7: Charge density difference map for a N adatom on a (5x5) unit cell. The small (light blue colored) atoms are N and large (magenta colored) atoms are Ga. The range of charge density (in e/Å^3) is shown in the bottom panel.

S11. (0,0') Ga-N adatom pair

A special case of (0,0') of Ga-N adatom pair, shown in Fig. 1 of the manuscript, is explained in some detail below. We observed that during geometry optimization for this pair of adatoms (Ga at the hcp site and N at the nearest fcc site before geometry optimization), the Ga adatom leaves its hcp site and goes nearly on top to the N adatom at fcc site. It is observed in both full-relax and fixed-slab calculations. Although this configuration is less stable than the (0,1') configuration of Ga-N pair (Fig. 1 of the manuscript). Here, we are showing the geometries of a such a Ga-N pair on a (5x5) and on a (3x3) unit cell before and after geometry optimization. These are shown in Figs. S8 and S9.

We started the calculations with N adatom at an fcc site and Ga adatom at a neighboring hcp site. The input configuration is shown in Fig. S8(a) and Fig. S9(a). Then we relaxed all the coordinates of all adatoms and top 2 GaN bilayers. The final configuration is shown in Fig. S8(b) and Fig. S9(b). We found that Ga adatom moves away from the hcp site and goes towards the N adatom.

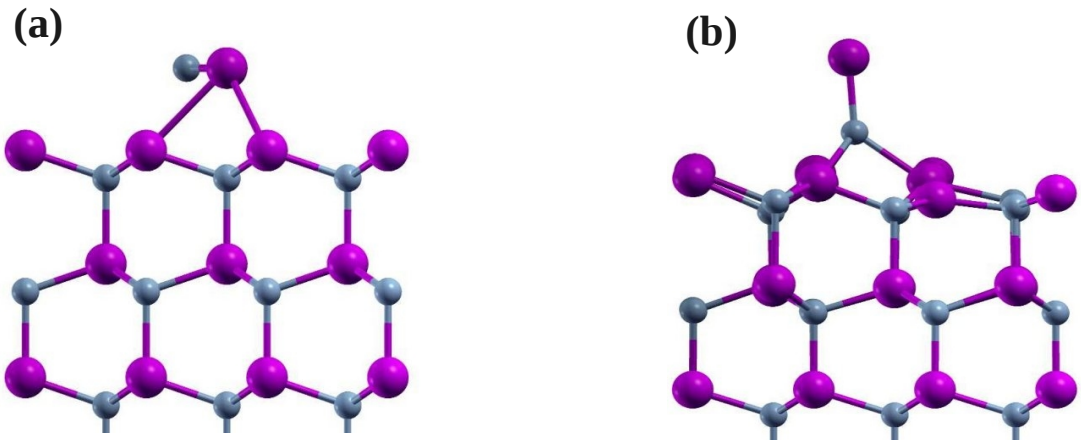

Fig. S8: Ga-N pair on (3×3) unit cell (a) before and (b) after geometry optimization.

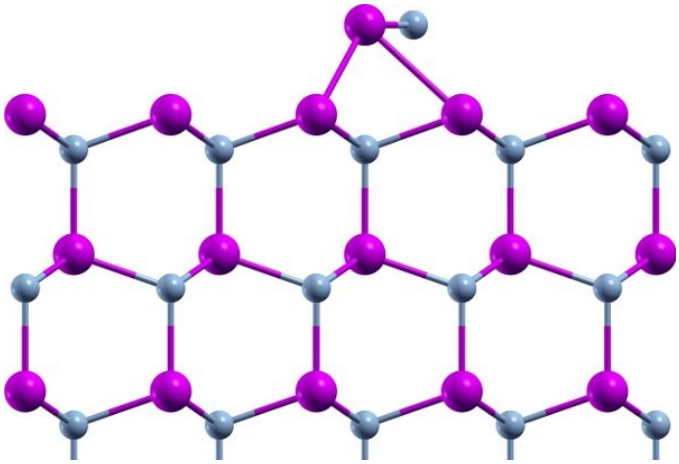

Fig. S9(a): Ga-N (0,0') pair on a (5x5) unit cell before geometry optimization.

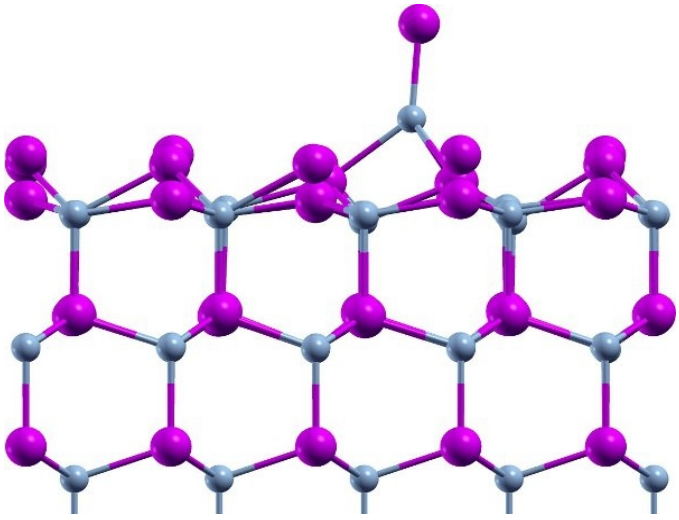

Fig. S9(b): Ga-N (0,0') pair on a (5x5) unit cell after geometry optimization.

S12. kinetic Monte Carlo simulations

We have developed a lattice-based kinetic Monte Carlo (kMC) simulation to model the homo-epitaxial growth of GaN(0001) surface. We have used a solid-on-solid (SOS) model to mimic the wurtzite GaN crystal, in which a Ga or N atom can be deposited (adsorbed) to a lattice site only if there is already a N or Ga atom present at that site. This procedure ensures that there would not be any vacancy in the growing GaN(0001) crystal. We have considered only diffusion and deposition of atoms in our kMC simulation and neglected evaporation, migration and interdiffusion processes. Each atom present at the surface is allowed to diffuse with a certain diffusion barrier. We have calculated these diffusion barriers with the help of DFT calculations, as has been explained in the manuscript.
For the kMC simulation result shown in Fig. 10 of the manuscript, we have employed the following simulation parameters.
System size = 128x128 with periodic boundary conditions
Temperature = 773 K
Ga : N flux ratio = 1
Coverage = 0.3 BLs (2457 atoms)
Flux Rate = 0.1 BL/s

S13. Relative surface energy calculations

GaN is a polar material, having alternate Ga and N layers along [0001] direction. It is not possible to obtain two equivalent surfaces by cutting the crystal perpendicular to [0001] direction. Hence, the absolute surface energy for a configuration on GaN(0001) surface is not possible to calculate. Instead, its relative surface energy with respect to some reference surface is typically calculated.

We have calculated the relative surface energies of few GaN surface configurations (with respect to clean Ga-terminated surface) and plotted them as a function of N chemical potential relative to nitrogen gas. The Ga-hcp and N-fcc configurations refer to the (2x2)-Ga adatom and (2x2)-N adatom configurations respectively. Ga-vacancy structure is obtained by removing one Ga atom from the topmost substrate Ga layer of a (2x2) unit cell and performing the geometry optimization of this configuration. N-vacancy structure is obtained by removing a N atom form the sub-surface N layer for a (2x2) unit cell and performing geometry optimization of this configuration. These relative surface energies are plotted in Fig. S10.

The right side of the x-axis corresponds to N-rich (Ga-poor) conditions where the chemical potential equals that of gaseous nitrogen. The left side of the x-axis corresponds to N-poor (Ga-rich) conditions where the chemical potential difference equals the formation energy of GaN crystal (~-1.00 eV). The allowed range of chemical potential of nitrogen has to be between GaN formation energy and bulk Nitrogen gas.

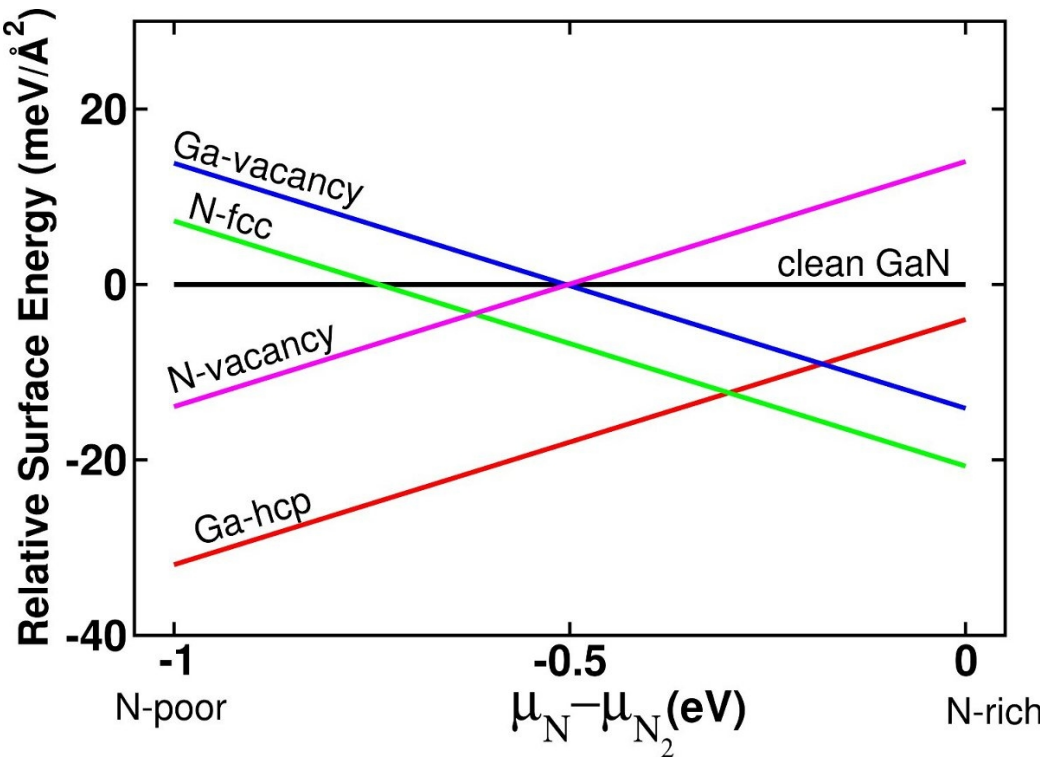

Fig. S10: Relative surface energy per unit area of various (2x2) GaN(0001) configurations. Clean GaN means Ga-terminated GaN(0001) surface, without any adsorbate. Gahcp(Nfcc) means single Ga(N) adatom at hcp(fcc) site on a (2x2) unit cell. Ga-vacancy(N-vacancy) means a Ga(N) atom is missing from the top GaN bilayer.

It is seen from these relative surface energy plots that under Ga-rich or N-poor conditions, the (2x2) Ga adatom configuration is most stable and under higher N-rich conditions, the (2x2) N adatom configuration is most favourable. The Ga-vacancy configuration is less stable than both the clean GaN surface and the (2x2) Ga-adatom configuration under Ga-rich conditions. It is only under extremely N-rich conditions, that the Ga-vacancy configuration becomes more favourable than Ga-adatom configuration. N-vacancy structure is stable under Ga-rich conditions. The above results are consistent with earlier work[5,6].

S14. (6x6) pairs calculations

We have also calculated the interaction energies between the adatoms on a (6x6) unit cell. We have only calculated for Ga-Ga and Ga-N pairs. We did this to see the effect of periodicity on the interaction energy values. Below we show the interaction energies calculated for the full-relax calculations for a (6x6) unit cell.

If we compare the interaction energies between adatoms on a (6x6) unit cell and on a (5x5) unit cell (Fig. 2 of the manuscript), we can say that for the same intersite distance, the magnitude of the

interaction energy differs. Since these calculations are done using FCS as the clean slab, they do not give the full picture. We see, like in Fig. 2 of the manuscript, that the interaction energies do not go to zero for large distances.
Because of large lattice distortions present at the GaN(0001) surface due to adsorbed species, we cannot compare the (5x5) and (6x6) calculations using FCS as the clean slab. We need to take DCS-rlx configuration as a reference for the clean slab for (6x6) like we had done for (5x5).

The periodic effects appear to change the trend (slightly) for the largest distance.

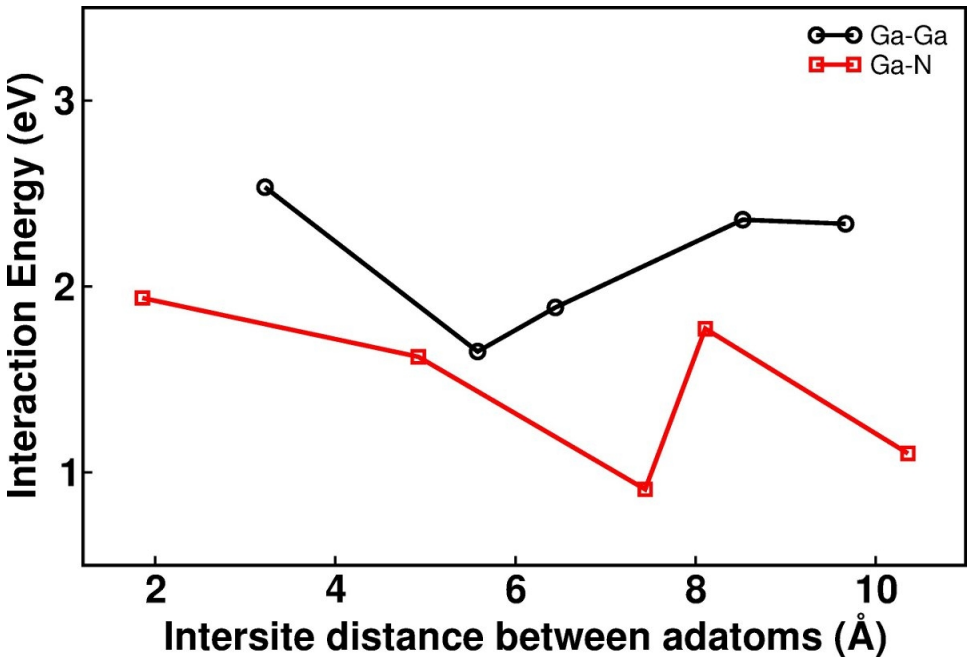

Fig. S11: Interaction energies as a function of intersite distances for the full-relaxed adatom pairs configurations on a (6x6) unit cell. Color code is shown in the figure.

S15. Sample input and output files

We show sample input and output files for the calculation mentioned in section S11.

INPUT FILE :

```
&CONTROL
   calculation = "relax",
   title = '3x3_UC',
   verbosity = 'high',
   restart_mode = 'from_scratch',
   prefix = 'GaN',
```

```
   pseudo_dir = '..../',
   outdir = '..../',
   tprnfor = .TRUE.,
   etot_conv_thr = 1.0D-04,
   forc_conv_thr = 1.0D-03,
   nstep = 400
 /
&SYSTEM
   ibrav = 0 ,
   celldm(1)= 6.088,
   nat = 83 ,
   ntyp = 3 ,
   ecutwfc = 70.0 ,
   ecutrho = 560.0 ,
   input_dft = 'PBE' ,
   occupations = 'smearing',
   smearing = 'methfessel-paxton' ,
   degauss = 0.05
 /

&ELECTRONS
  electron_maxstep = 500 ,
  diagonalization='cg',
  conv_thr = 1.D-7,
  mixing_mode = 'local-TF',
  mixing_beta = 0.5D0
/

&IONS
  ion_dynamics = 'bfgs'
/

CELL_PARAMETERS {hexagonal}
 3.00  0.00   0.00
 1.50  2.598  0.00
 0.00  0.00   9.756

ATOMIC_SPECIES
  Ga  69.723  Ga.pbe-n-van.UPF
  N   14.006  N.pbe-van_ak.UPF
  H   1.0079  H.pz-vbc_075.UPF

ATOMIC_POSITIONS {crystal}
H       0.000000000   0.000000000   0.196479930   0   0   0
H       0.333333333   0.000000000   0.196479930   0   0   0
H       0.666666666   0.000000000   0.196479930   0   0   0
H       0.000000000   0.333333333   0.196479930   0   0   0
H       0.000000000   0.666666666   0.196479930   0   0   0
H       0.333333333   0.333333333   0.196479930   0   0   0
H       0.333333333   0.666666666   0.196479930   0   0   0
```

H	0.666666666	0.333333333	0.196479930	0	0	0
H	0.666666666	0.666666666	0.196479930	0	0	0
N	0.000000000	0.000000000	0.229502345	0	0	0
N	0.333333333	0.000000000	0.229502345	0	0	0
N	0.666666666	0.000000000	0.229502345	0	0	0
N	0.000000000	0.333333333	0.229502345	0	0	0
N	0.000000000	0.666666666	0.229502345	0	0	0
N	0.333333333	0.333333333	0.229502345	0	0	0
N	0.333333333	0.666666666	0.229502345	0	0	0
N	0.666666666	0.333333333	0.229502345	0	0	0
N	0.666666666	0.666666666	0.229502345	0	0	0
Ga	0.111111111	0.111111111	0.249992062	0	0	0
Ga	0.111111111	0.444444444	0.249992062	0	0	0
Ga	0.111111111	0.777777777	0.249992062	0	0	0
Ga	0.444444444	0.111111111	0.249992062	0	0	0
Ga	0.444444444	0.444444444	0.249992062	0	0	0
Ga	0.444444444	0.777777777	0.249992062	0	0	0
Ga	0.777777777	0.111111111	0.249992062	0	0	0
Ga	0.777777777	0.444444444	0.249992062	0	0	0
Ga	0.777777777	0.777777777	0.249992062	0	0	0
N	0.111111111	0.111111111	0.312865782	0	0	0
N	0.111111111	0.444444444	0.312865782	0	0	0
N	0.111111111	0.777777777	0.312865782	0	0	0
N	0.444444444	0.111111111	0.312865782	0	0	0
N	0.444444444	0.444444444	0.312865782	0	0	0
N	0.444444444	0.777777777	0.312865782	0	0	0
N	0.777777777	0.111111111	0.312865782	0	0	0
N	0.777777777	0.444444444	0.312865782	0	0	0
N	0.777777777	0.777777777	0.312865782	0	0	0
Ga	0.000000000	0.000000000	0.333276902	0	0	0
Ga	0.333333333	0.000000000	0.333276902	0	0	0
Ga	0.666666666	0.000000000	0.333276902	0	0	0
Ga	0.000000000	0.333333333	0.333276902	0	0	0
Ga	0.000000000	0.666666666	0.333276902	0	0	0
Ga	0.333333333	0.333333333	0.333276902	0	0	0
Ga	0.333333333	0.666666666	0.333276902	0	0	0
Ga	0.666666666	0.333333333	0.333276902	0	0	0
Ga	0.666666666	0.666666666	0.333276902	0	0	0
N	0.000000000	0.000000000	0.396156112			
N	0.333333333	0.000000000	0.396156112			
N	0.666666666	0.000000000	0.396156112			
N	0.000000000	0.333333333	0.396156112			
N	0.000000000	0.666666666	0.396156112			
N	0.333333333	0.333333333	0.396156112			
N	0.333333333	0.666666666	0.396156112			
N	0.666666666	0.333333333	0.396156112			
N	0.666666666	0.666666666	0.396156112			
Ga	0.111111111	0.111111111	0.416443231			
Ga	0.111111111	0.444444444	0.416443231			
Ga	0.111111111	0.777777777	0.416443231			

```
Ga      0.444444444   0.111111111   0.416443231
Ga      0.444444444   0.444444444   0.416443231
Ga      0.444444444   0.777777777   0.416443231
Ga      0.777777777   0.111111111   0.416443231
Ga      0.777777777   0.444444444   0.416443231
Ga      0.777777777   0.777777777   0.416443231
N       0.111111111   0.111111111   0.479347469
N       0.111111111   0.444444444   0.479347469
N       0.111111111   0.777777777   0.479347469
N       0.444444444   0.111111111   0.479347469
N       0.444444444   0.444444444   0.479347469
N       0.444444444   0.777777777   0.479347469
N       0.777777777   0.111111111   0.479347469
N       0.777777777   0.444444444   0.479347469
N       0.777777777   0.777777777   0.479347469
Ga      0.000000000   0.000000000   0.499538888
Ga      0.333333333   0.000000000   0.499538888
Ga      0.666666666   0.000000000   0.499538888
Ga      0.000000000   0.333333333   0.499538888
Ga      0.000000000   0.666666666   0.499538888
Ga      0.333333333   0.333333333   0.499538888
Ga      0.333333333   0.666666666   0.499538888
Ga      0.666666666   0.333333333   0.499538888
Ga      0.666666666   0.666666666   0.499538888
Ga      0.444444444   0.444444444   0.560000000
N       0.555555555   0.555555555   0.560000000

K_POINTS (automatic)
3  3  1  0  0  0
```

OUTPUT FILE :

```
End of BFGS Geometry Optimization
    Final energy   =   -7327.3059057830 Ry
Begin final coordinates

ATOMIC_POSITIONS (crystal)
H       0.000000000   0.000000000   0.196479930   0   0   0
H       0.333333333   0.000000000   0.196479930   0   0   0
H       0.666666666   0.000000000   0.196479930   0   0   0
H       0.000000000   0.333333333   0.196479930   0   0   0
H       0.000000000   0.666666666   0.196479930   0   0   0
H       0.333333333   0.333333333   0.196479930   0   0   0
H       0.333333333   0.666666666   0.196479930   0   0   0
H       0.666666666   0.333333333   0.196479930   0   0   0
H       0.666666666   0.666666666   0.196479930   0   0   0
N       0.000000000   0.000000000   0.229502345   0   0   0
```

N	0.333333333	0.000000000	0.229502345	0	0	0
N	0.666666666	0.000000000	0.229502345	0	0	0
N	0.000000000	0.333333333	0.229502345	0	0	0
N	0.000000000	0.666666666	0.229502345	0	0	0
N	0.333333333	0.333333333	0.229502345	0	0	0
N	0.333333333	0.666666666	0.229502345	0	0	0
N	0.666666666	0.333333333	0.229502345	0	0	0
N	0.666666666	0.666666666	0.229502345	0	0	0
Ga	0.111111111	0.111111111	0.249992062	0	0	0
Ga	0.111111111	0.444444444	0.249992062	0	0	0
Ga	0.111111111	0.777777777	0.249992062	0	0	0
Ga	0.444444444	0.111111111	0.249992062	0	0	0
Ga	0.444444444	0.444444444	0.249992062	0	0	0
Ga	0.444444444	0.777777777	0.249992062	0	0	0
Ga	0.777777777	0.111111111	0.249992062	0	0	0
Ga	0.777777777	0.444444444	0.249992062	0	0	0
Ga	0.777777777	0.777777777	0.249992062	0	0	0
N	0.111111111	0.111111111	0.312865782	0	0	0
N	0.111111111	0.444444444	0.312865782	0	0	0
N	0.111111111	0.777777777	0.312865782	0	0	0
N	0.444444444	0.111111111	0.312865782	0	0	0
N	0.444444444	0.444444444	0.312865782	0	0	0
N	0.444444444	0.777777777	0.312865782	0	0	0
N	0.777777777	0.111111111	0.312865782	0	0	0
N	0.777777777	0.444444444	0.312865782	0	0	0
N	0.777777777	0.777777777	0.312865782	0	0	0
Ga	0.000000000	0.000000000	0.333276902	0	0	0
Ga	0.333333333	0.000000000	0.333276902	0	0	0
Ga	0.666666666	0.000000000	0.333276902	0	0	0
Ga	0.000000000	0.333333333	0.333276902	0	0	0
Ga	0.000000000	0.666666666	0.333276902	0	0	0
Ga	0.333333333	0.333333333	0.333276902	0	0	0
Ga	0.333333333	0.666666666	0.333276902	0	0	0
Ga	0.666666666	0.333333333	0.333276902	0	0	0
Ga	0.666666666	0.666666666	0.333276902	0	0	0
N	-0.000535485	-0.000478220	0.398047665			
N	0.332307890	0.001641478	0.395284613			
N	0.669279705	-0.001992202	0.396768575			
N	0.001576562	0.332401466	0.395244027			
N	-0.001990396	0.669245591	0.396804607			
N	0.332921479	0.333014499	0.395736856			
N	0.331959038	0.667846271	0.395792011			
N	0.667885195	0.331921092	0.395769646			
N	0.666942555	0.666934110	0.395585427			
Ga	0.112615969	0.113064286	0.417459922			
Ga	0.107183685	0.446998947	0.416030253			
Ga	0.113213252	0.774167068	0.418638629			
Ga	0.447159926	0.107325830	0.415938837			
Ga	0.445364207	0.445216101	0.414398118			
Ga	0.443740265	0.780234037	0.415616401			

Ga 0.774022867 0.113251473 0.418754304
Ga 0.780390022 0.443606102 0.415754671
Ga 0.776226453 0.776415929 0.419632287
N 0.112629897 0.113089063 0.481426714
N 0.109323225 0.446463550 0.479096347
N 0.118234628 0.772155624 0.485537131
N 0.446661685 0.109417141 0.479044332
N 0.442663030 0.442502555 0.476971391
N 0.440518304 0.785112770 0.476845077
N 0.772115741 0.118245571 0.485419281
N 0.785282364 0.440423283 0.476847259
N 0.772360721 0.772532356 0.483425205
Ga 0.000493116 0.000976887 0.492105844
Ga 0.337737577 -0.000968809 0.504220514
Ga 0.662458866 -0.005373210 0.509122421
Ga -0.001802089 0.338549794 0.506055672
Ga -0.005377501 0.662393608 0.507829113
Ga 0.331637950 0.331773497 0.487062128
Ga 0.355356305 0.658120220 0.501403325
Ga 0.658640399 0.355043717 0.501295302
Ga 0.656456372 0.655832913 0.500333872
Ga 0.573144368 0.570768637 0.601892745
N 0.554277143 0.554738410 0.541425120
End final coordinates

References

1. G. Henkelman, A. Arnaldsson and H. Jónsson, *Comput. Mater. Sci.*, 2006, **36**, 354-360.
2. E. Sanville, S. D. Kenny, R. Smith and G. Henkelman, *J. Comp. Chem.*, 2007, **28**, 899-908.
3. P. L. Silverstrelli, A. Ambrosetti, S. Grubisiĉ and F. Ancilotto, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2012, **85**, 165405.
4. X. Liu, C.-Z. Wang, H.-Q. Lin, K. Chang, J. Chen and K.-M. Ho, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2015, **91**, 035415.
5. A. L. Rosa and J. Neugebauer, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2006, **73**, 205346.
6. J. Neugebauer and C. G. Walle, In *Advances in Solid State Physics 35*; Helbig, R., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1995; Chapter- Native defects and impurities in GaN, pp 25-44.