Supporting Information

Towards Understanding of Gold Interaction with AIII-BV Semiconductors at Atomic Level

B.R. Jany[a*], A. Janas[a], W. Piskorz[c], K. Szajna[a], A. Kryshtal[b], G. Cempura[b], P. Indyka[c,d], A. Kruk[b], A. Czyrska-Filemonowicz[b], F. Krok[a]

[a]The Marian Smoluchowski Institute of Physics, Jagiellonian University, Lojasiewicza 11, 30-348 Krakow, Poland
[b]The International Centre of Electron Microscopy for Materials Science, AGH University of Science and Technology, 30-059 Krakow, Poland
[c]The Faculty of Chemistry, Jagiellonian University, ul. Gronostajowa 2, 30-387 Krakow, Poland
[d]Malopolska Centre of Biotechnology (MCB), Jagiellonian University, Krakow, 30-387, Poland

*Corresponding author e-mail: benedykt.jany@uj.edu.pl

Machine Learning HAADF STEM image quantification

The HAADF STEM imaging mode provides structural images where intensities are proportional to both the thickness and mean atomic number Z[1] (Au columns appear brighter than In-Sb columns). By assuming a constant sample thickness, this can be efficiently used to distinguish between columns containing Au atoms and pure In-Sb columns. The image was segmented into cells containing atomic columns[2,3]. This has the advantage that all image scattering is associated to some atomic column. Due to the noise inherent to the experiment, intermixing between the Au and In-Sb phases is studied atomic column-by-column, similar as in[4]. Here we used Machine Learning algorithms (Random Forest) as implemented in Trainable Weka Segmentation[5] to statistically distinguish between these phases, as successively used in HAADF STEM Tomography[6]. For the Trainable Weka Segmentation the following training features were used: mean, median, variance, maximum, minimum. The classes were balanced. The rest settings were set on their default values (classifier: fast random forest of 200 trees with 2 features per tree). The reference areas were used for the Au atomic columns and In-Sb atomic columns, as indicated in Fig. S1a. As a result of image quantification of analysis area (area with Au atom diffusion into the bulk InSb crystal -Fig. S1a) the Au probability map and InSb probability map is computed Fig. S1b-c. Since the probability of finding Au is directly proportional to the number of Au atoms in the atomic row (assuming constant sample thickness) and $Probability_{Au}+Probability_{InSb}=1$, so the probability of finding Au is directly the Au atoms concentration in the sample (value “1” corresponds to the concentration of 100 atomic %). The histogram of Au probability (Au atomic concentration) from the area with Au atom diffusion into the bulk InSb crystal is presented in Fig. S1d. Several local maxima are seen Fig. S1d and Table S1. This is compared with EDX measurements of this region as presented in histogram of Au atomic concentration from EDX Fig. 1e and Table S1. The quantitative HAADF STEM shows more local maxima in comparison to the EDX measurements, so in this case in more locally sensitive to the changes of Au atomic concentrations.

HAADF STEM quantification Au atomic concentrations [X at.]	**EDX measurements** Au atomic concentration [X at.]
0.037	–
0.075	0.068
0.12	–
0.23	–
0.28	–

Table S1: Au atomic concentaration in the area with Au atom diffusion into the bulk InSb crystal as measured by HAADF STEM image quantification and EDX. In this case the quantitative HAADF measurements are more sensitive to local Au concentration then EDX, since more local maxima are visible (value “1” corresponds to the concentration of 100 atomic %).

To check the validity of whole chemical quantification procedure i.e. how it estimates the true unknown Au concentration, we compared the results of HAADF STEM quantification with EDX measurements. Assuming that EDX concentration (0.068) is a mixture of two components as seen by HAADF quantification (0.037 and 0.075), which gives central value of 0.060, one can estimate the maximal relative quantification uncertainty as relative difference (0.068-0.060)/0.060*100% which is equal to 12.6%.

In Fig. S1f Au probability map (Au atomic concentration) from Area1 is presented, one can see in details that different amount of Au atoms are built up into the In-Sb lattice atomic positions.

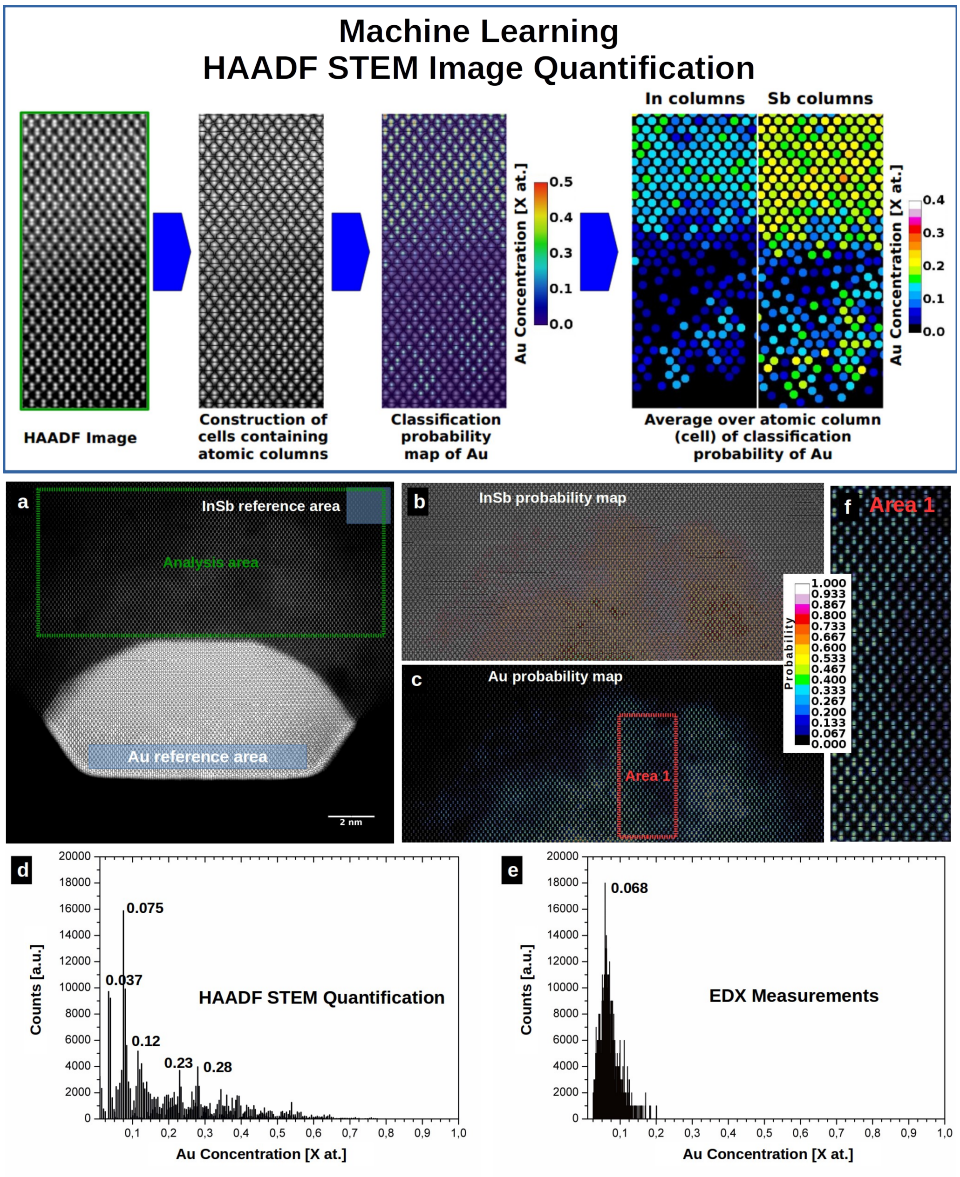

Fig. S1: HAADF STEM image quantification of Au/InSb at 330C. Graphical presentation of the idea of the method (top). a) atomically resolved HAADF STEM image, Au and InSb reference area marked, b) quantification result, InSb probability map overlaid on HAADF image from analysis area, c) quantification result, Au probability map overlaid on HAADF image from analysis area, d) histogram of Au probability from analysis area. Since the probability of finding Au is directly proportional to the number of Au atoms in the atomic row (assuming constant sample thickness) and ProbabilityAu+ProbabilityInSb=1, so the probability of finding Au is directly the Au atoms concentration in the sample. e) histogram of Au atomic concentration, as measured by EDX. The obtained Au probability (Au atomic concentration) from HAADF STEM quantification matches well the EDX measurements. The main maxima, corresponding to the main atomic concentrations are marked. It is seen that HAADF STEM is more locally sensitive than EDX since additional small maxima are visible. f) Au probability map (Au atomic concentration) from Area1, one can see in details that different amount of Au atoms are built up into the In-Sb lattice atomic positions. Value "1" corresponds to the concentration of 100 atomic %.

Local sample thickness variations changes for the HAADF quantification area of the thin foil sample prepared by FIB were estimated, and their influence on the chemical quantification result.

The line profile 1 (Fig. S1-1b) was extracted from HAADF STEM image (Fig. S1-1a), from the area without Au diffusion. Since there is only one material in this area (namely InSb) so the changes of the HAADF Intensity in this area are only due to the local sample thickness changes, as seen in the line profile. Taking the histogram of the intensities from the linear profile Fig. S1-1c, one can calculate the mean HAADF intensity and the standard deviation of it. So the relative HAADF signal deviation related to the local sample thickness changes is equal to 0.28%.

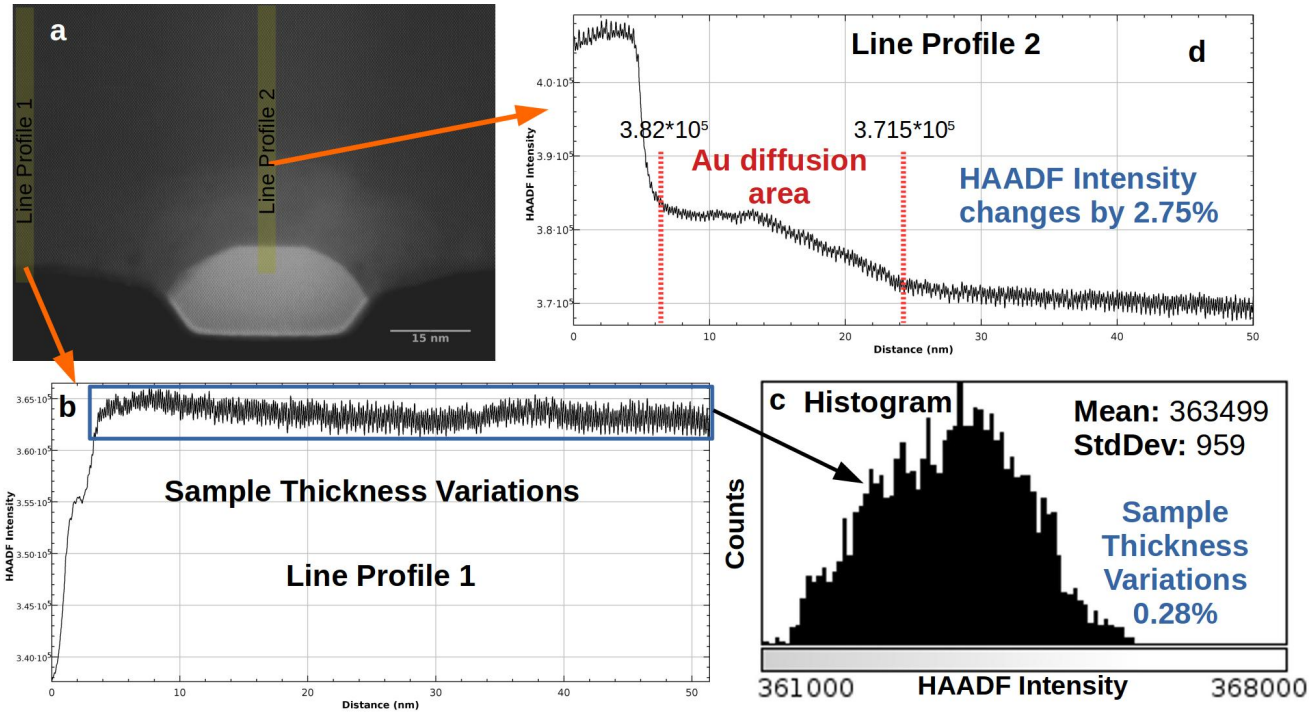

*Fig. S1-1: Local sample thickness variation changes for the HAADF quantification area of the thin foil sample of Au/InSb at 330C. a) HAADF STEM image of the analysis area, two line profiles indicated. b) Line profile 1 from HAADF STEM image a) from sample area without Au diffusion, the local HAADF signal changes correspond to local sample thickness variations. c) Histogram from Line profile 1 b) from the marked by blue rectangle area on the profile. It is seen that the distribution standard deviation (value of 958) is equal to 0.28% of the mean value (363499). So the relative HAADF signal deviation related to the local sample thickness changes is equal to 0.28%. d) Line profile 2 from HAADF STEM image a) through the nanowire and Au diffusion area. It is seen that the HAADF signal changes on the Au diffusion area by 2.75% (changes from 3.82*10^5 to 3.715*10^5) are above deviations related to the estimated local sample thickness changes 0.28%.*

One can now compare this values with the HAADF signal changes below the nanowire in the Au diffusion area, see Fig. S1-1d. It is seen that here the HAADF signal changes by 2.75% (changes from 3.82*10^5 to 3.715*10^5), which is related to Au diffusion into InSb. This HAADF signal changes related to the Au diffusion are above deviations related to the estimated local sample thickness changes. Next, one can estimate the influence of the local sample thickness changes on the HAADF chemical quantification performed by comparing the estimated values 0.28%/2.75%*100%=10.2%. So the estimated relative uncertainty of the performed HAADF STEM chemical quantification due to the local sample thickness variation is 10.2%.

Au-AIII-BV Phase Diagrams

The Au-AIII-BV phase diagrams were calculated from First Principles using the generalized gradient approximation (GGA) approximation to density functional theory (DFT) and the DFT+U extension to it[7,8] by the Materials Project[9] and also OQMD[10,11]. In agreement with experimental phase diagrams[12–14].

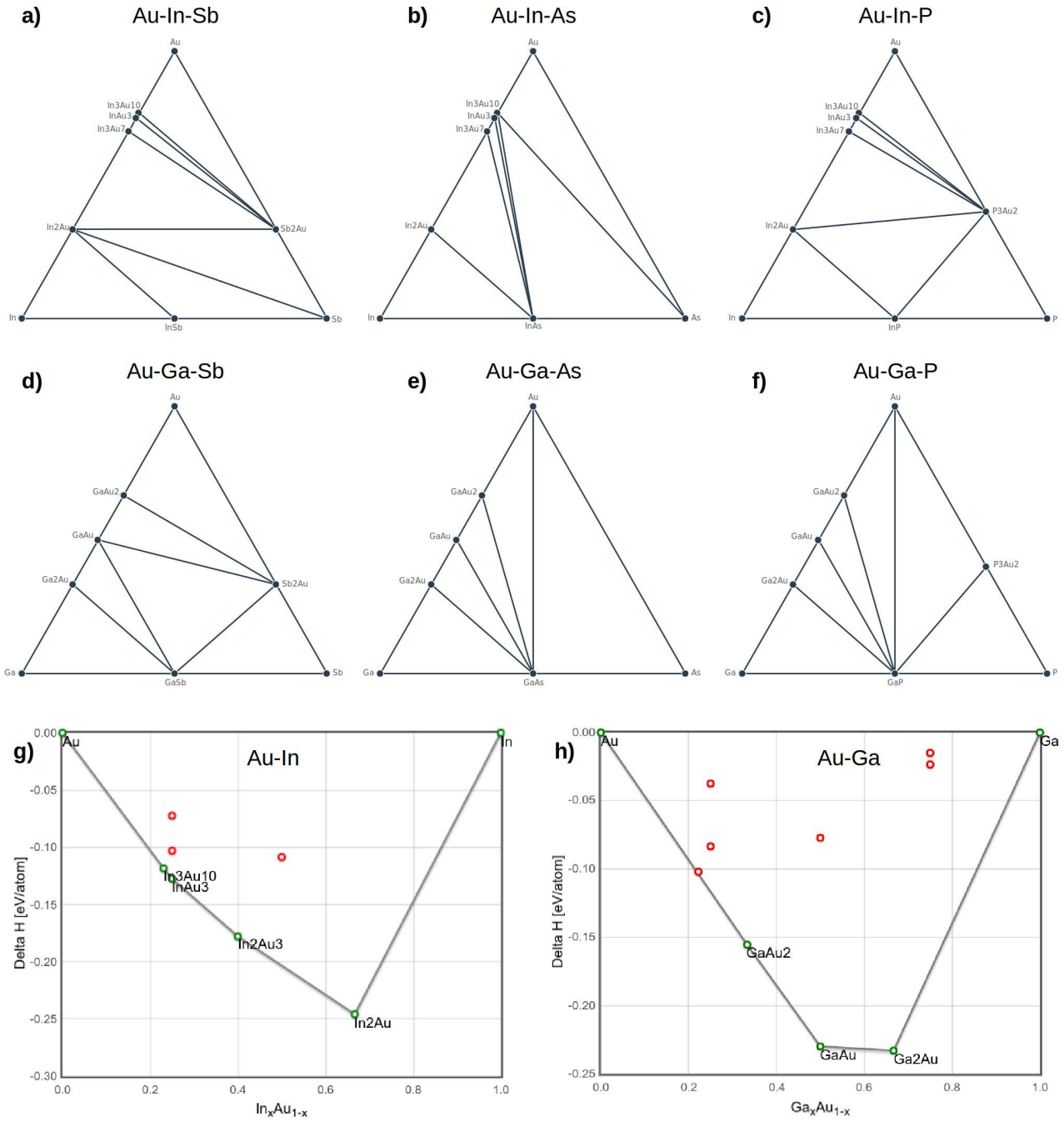

Fig. S2: Theoretically calculated Phase Diagrams for Au-AIII-BV systems: a) Au-In-Sb, b) Au-In-As, c) Au-In-P, d) Au-Ga-Sb, e) Au-Ga-As, f) Au-Ga-P, g) Au-In, h)Au-Ga by the Materials Project[9] and also OQMD[10].

Au-In System		**Au-Ga System**	
Phase	Formation Energy [eV]	Phase	Formation Energy [eV]
$AuIn_2$	-0.246	$AuGa_2$	-0.234
Au_3In	-0.126	AuGa	-0.227
Au_7In_3	-0.141	Au_2Ga	-0.151
$Au_{10}In_3$	-0.116	Au_7Ga_2	-0.102
AuIn	-0.108	Au_3Ga	-0.083
$AuIn_3$	0.016	$AuGa_3$	-0.023

Table S2: Theoretically calculated Phases in the Au-In and Au-Ga system together with their formation energies by the Materials Project[9] *and also OQMD*[10]*. Unstable phases also included. In agreement with experimental phase diagrams*[15,16]*.*

RHEED patterns of the atomically clean and reconstructed AIII-BV surfaces

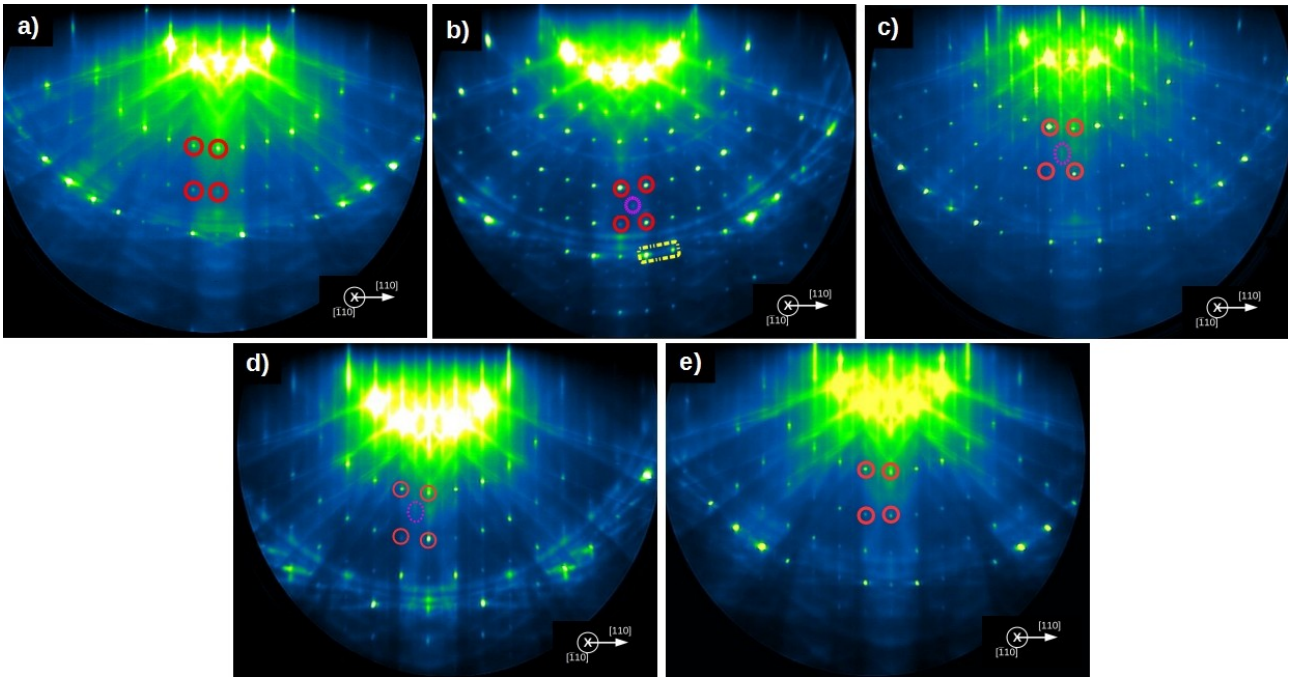

Fig. S3: RHEED patterns of the atomically clean and reconstructed surfaces of a) (3x1) GaSb(001), b) c(8x2) GaAs(001), c) (4x2) GaP(001), d) (4x2) InAs(001), e) (4x2) InP(001). Main reconstruction spots marked.

Multivariate Statistical Analysis

The obtained from SEM, AFM and TEM measurements data on nanostructures formed in the Au/AIIIBV systems are presented in Table S3, as in detailed described in the main article. It is seen that for each AIIIBV system a set of seven parameters is used to describe it, forming seven dimensional parameter space. This is later used for multivariate statistical analysis using machine learning Multidimensional Scaling (MDS)[17] and k-means clustering[18].

AIIIBV system	Average size [nm]	Surface density [1/um^2]	Surface diffusion radius [nm]	Average height [nm]	Nanostructure percent under the sample surface [%]	Au concentration in the nanostructure [atomic %]	Number of Au atoms needed to release one AIII metallic atom on the surface
InSb	146.3	6.1	202.3	8.0	35	33.3	0.5
InAs	35.9	170.5	38.3	7.9	0	75.0	3
InP	31.4	214.4	34.1	11.3	6.45	33.3	0.5
GaSb	13.1	747.4	18.3	0.9	69	33.3	0.5
GaAs	8.2	2003.8	11.2	1.1	41.8	100.0	2.5
GaP	5.8	2088.5	10.9	1.4	13.4	100.0	5.9

Table S3: Measured seven parameters of the formed nanostructures in the Au/AIIIBV systems.

Formed Nanostructures Properties

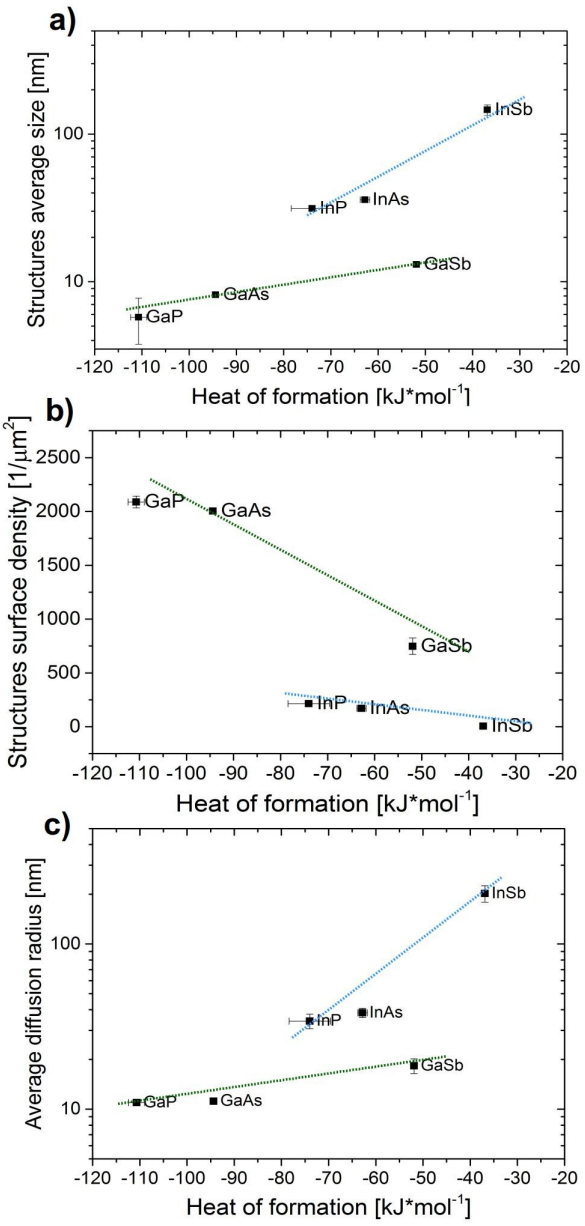

Fig. S4: Nanostructures resulted from 2ML Au deposited on InSb(001), GaSb(001), GaAs(001), InP(001) and InAs(001) surfaces at 330C. Average sizes b), surface densities c) and average diffusion radius d) as a function of the heat of formation of AIII-BV semiconductor. The lines are drawn to guide the eye. Different behavior is observed for the structures grown on In and Ga rich surfaces.

Results of STEM EDX Analysis

For the examined Au/AIII-BV systems the STEM EDX measurements were performed in form of hyperspectral maps. To obtain EDX signal coming only from formed nanostructures the Machine Learning Blind Source Separation (BSS) analysis using Non Negative Matrix Factorization (NMF) was performed on hyperspectral EDX data accordingly to [19,20]. Later the chemical compositions were quantified by Cliff-Lorimer method, the results of quantification are presented in Table S4.

Au/AIII-BV System	Results of EDX of bulk AIII-BV [at. %]	Results of EDX of Nanostructures [at. %]
2ML Au on InSb(001)	In: 51.0(1.0) Sb: 49.0(1.0)	Au: 34.0(1.0) In: 66.0(1.0)
2ML Au on InAs(001)	In: 51.3(1.3) As: 48.7(1.3)	Au: 73.3(1.3) In: 26.7(1.3)
2ML Au on InP(001)	In: 47.3(2.7) P:52.7(2.7)	Au: 28.6(2.7) In: 71.4(2.7)
2ML Au on GaSb(001)	Ga: 54.1(4.1) Sb:45.9(4.1)	Au: 37.1(4.1) Ga: 62.9(4.1)
2ML Au on GaAs(001)	Ga: 47.9(2.1) As: 52.1(2.1)	Pure Au
2ML Au on GaP(001)	Ga: 52.7(2.7) P: 47.1(2.7)	Pure Au

Table S4: Results of STEM EDX analysis for the Au/AIII-BV systems.

References

(1) Krivanek, O. L.; Chisholm, M. F.; Nicolosi, V.; Pennycook, T. J.; Corbin, G. J.; Dellby, N.; Murfitt, M. F.; Own, C. S.; Szilagyi, Z. S.; Oxley, M. P.; et al. Atom-by-Atom Structural and Chemical Analysis by Annular Dark-Field Electron Microscopy. *Nature* **2010**, *464* (7288), 571–574.

(2) Rosenauer, A.; Mehrtens, T.; Müller, K.; Gries, K.; Schowalter, M.; Satyam, P. V.; Bley, S.; Tessarek, C.; Hommel, D.; Sebald, K.; et al. Composition Mapping in InGaN by Scanning Transmission Electron Microscopy. *Ultramicroscopy* **2011**, *111* (8), 1316–1327. http://dx.doi.org/10.1016/j.ultramic.2011.04.009.

(3) Nguyen, D. T.; Findlay, S. D.; Etheridge, J. The Spatial Coherence Function in Scanning Transmission Electron Microscopy and Spectroscopy. *Ultramicroscopy* **2014**, *146*, 6–16. http://dx.doi.org/10.1016/j.ultramic.2014.04.008.

(4) Jany, B. R.; Gauquelin, N.; Willhammar, T.; Nikiel, M.; Van Den Bos, K. H. W.; Janas, A.; Szajna, K.; Verbeeck, J.; Van Aert, S.; Van Tendeloo, G.; et al. Controlled Growth of Hexagonal Gold Nanostructures during Thermally Induced Self-Assembling on Ge(001) Surface. *Sci. Rep.* **2017**, *7*.

(5) Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I. H. The WEKA Data Mining Software: An Update. *SIGKDD Explor Newsl* **2009**, *11* (1), 10–18. https://doi.org/10.1145/1656274.1656278.

(6) Staniewicz, L.; Vaudey, T.; Degrandcourt, C.; Couty, M.; Gaboriaud, F.; Midgley, P. Electron Tomography Provides a Direct Link between the Payne Effect and the Inter-Particle Spacing of Rubber Composites. *Sci. Rep.* **2014**, *4*, 7389–.

(7) Ong, S. P.; Wang, L.; Kang, B.; Ceder, G. Li-Fe-P-O2 Phase Diagram from First Principles Calculations. *Chem. Mater.* **2008**, *20* (5), 1798–1807. https://doi.org/10.1021/cm702327g.

(8) Jain, A.; Hautier, G.; Ong, S.; Moore, C.; Fischer, C.; Persson, K.; Ceder, G. Formation Enthalpies by Mixing GGA and GGA + U Calculations. *Phys. Rev. B* **2011**, *84* (4), 045115. https://doi.org/10.1103/PhysRevB.84.045115.

(9) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; et al. The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1* (1), 011002. https://doi.org/10.1063/1.4812323.

(10) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). *JOM* **2013**, *65* (11), 1501–1509. https://doi.org/10.1007/s11837-013-0755-4.

(11) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies. *Npj Comput. Math.* **2015**, *1*, 15010. https://doi.org/10.1038/npjcompumats.2015.10.

(12) Thomas, T. C.; Williams, R. S. Solid Phase Equilibria in the Au-Ga-As, Au-Ga-Sb, Au-In-As, and Au-In-Sb Ternaries. *J. Mater. Res.* **4AD**, *1* (2), 352–360. https://doi.org/10.1557/JMR.1986.0352.

(13) Hassam, S.; Rogez, J.; Bahari, Z. Experimental Phase Diagram of the AuSb–InSb Section in the Au–In–Sb System. *J. Chem. Thermodyn.* **2014**, *70*, 168–175. http://dx.doi.org/10.1016/j.jct.2013.09.025.

(14) C.T.Tsai, J. H. P.; Williams, R. S. INTERFACIAL CHEMICAL REACTIONS OF Au WITH III–V COMPOUND SEMICONDUCTORS. *Off. Nav. Res. Tech. Rep.* **1988**, *AD-A198-610* (8), 1–23.

(15) Elliott, R. P.; Shunk, F. A. The Au-Ga (Gold-Gallium) System. *Bull. Alloy Phase Diagr.* **1981**, *2* (3), 356–358.

(16) Okamoto, H. Au-In (Gold-Indium). *J Phase Equilib Diffus* **2004**, *25* (2), 197–198.

(17) Torgerson, W. S. Multidimensional Scaling: I. Theory and Method. *Psychometrika* **1952**, *17* (4), 401–419.
(18) MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations. In *Fifth Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, Calif., 1967; pp 281–297.
(19) Rossouw, D.; Burdet, P.; de la Peña, F.; Ducati, C.; Knappett, B. R.; Wheatley, A. E. H.; Midgley, P. A. Multicomponent Signal Unmixing from Nanoheterostructures: Overcoming the Traditional Challenges of Nanoscale X-Ray Analysis via Machine Learning. *Nano Lett.* **2015**, *15* (4), 2716–2720. https://doi.org/10.1021/acs.nanolett.5b00449.
(20) Jany, B. R.; Janas, A.; Krok, F. Retrieving the Quantitative Chemical Information at Nanoscale from Scanning Electron Microscope Energy Dispersive X-Ray Measurements by Machine Learning. *Nano Lett.* **2017**, *17* (11), 6520–6525. https://doi.org/10.1021/acs.nanolett.7b01789.