4

Image Processing Quality Analysis for Particle Based Peptide Array Production on a Microchip

JennyWagner[1,2,6] et al[*]
[1]*Kirchhoff Institute for Physics, Heidelberg University*
[2]*German Cancer Research Centre*
[6]*Frankfurt Institute for Advanced Studies*
Germany

1. Introduction

1.1 Motivation

Highly complex microarray systems based on combinatorial synthesis techniques are in wide-spread use in biological, medical and pharmaceutical research Chee et al. (1996); Cretich et al. (2006); Debouck & Goodfellow (1999). Two prominent examples are micro arrays for the artificial synthesis of arbitrary DNA sequences out of nucleic acids Heller (2002) and peptide synthesis out of amino acids Beyer et al. (2007); Templin et al. (2003). In the case of DNA arrays, these experiments mostly focus on gene identification or gene expression profiling to determine the effects of single genes on cellular evolution. Peptide arrays aim at understanding interactions of peptides with other molecules. As sequences in proteins, peptides are involved in the regularisation of biological activity.

Since a large number of combinations that build a valid molecule chain is possible, both micro array variants require high densities in the range of 10,000 cDNAs or peptides per cm^2. Density and size of the synthesis sites are the critical parameter for such systems, since they limit the number of molecule sorts synthesized per array. Using the micro array in experiments of bio-medical research or in clinical diagnostics means bringing the array in touch with target molecules. Hence, the decreasing size of the synthesis sites, also reduces the amount of target molecules required (i.e. proteins or antibodies), which may be expensive or hard to obtain.

With increasing array complexity, the need for an automated read-out of the results and structuring of the acquired information arises because manual evaluation is tedious and often inconsistent when comparing thousands of synthesis sites with each other. Here, image

[*] Felix Löffler[1,2], Tobias Förtsch[1,2], Christopher Schirwitz[2], Simon Fernandez[2], Heinz Hinkers[3], Heinrich F. Arlinghaus[4], Florian Painke[1], Kai König[1,2,5], Ralf Bischoff[2], Alexander Nesterov-Müller[5], Frank Breitling[5], Michael Hausmann[1] and Volker Lindenstruth[6]
[1]*Kirchhoff Institute for Physics, Heidelberg University, Germany*
[2]*German Cancer Research Centre, Germany*
[3]*Verbundzentrum für Oberflächenanalyse Münster, Germany*
[4] *Institute of Physics, University of Muenster, Germany*
[5]*Karlsruhe Institute for Technology, Germany*
[6]*Frankfurt Institute for Advanced Studies, Germany*

acquisition and image processing based evaluation can be used. While the concept of DNA arrays includes image processing in many ways Angulo & Serra (2003); Bajcsy (2005); Brown et al. (2001), querying the literature yields only a few image processing supported peptide array experiments. In most cases, mass spectrometry data is analysed by means of pattern matching in order to identify peptide sequences for quality analysis Gusev et al. (1993).

Yet, up to now, all image processing tasks have only been used in the *evaluation* step of the array experiment or as a post-processing routine to compensate for irregularities in the quality of the array Wang et al. (2001), but not during array production itself. Contrary to that, we propose a destruction free quality control system based on image acquisition and image processing that is capable of monitoring the peptide array production process to improve the quality of the array with respect to the density of correctly assembled peptides. While post-processing may not always restore the information content of defective synthesis sites, our method detects insufficiencies and errors when they occur, so that subsequent corrections lead to fully functional arrays.

1.2 Outline of the chapter

Section 2 starts with describing the process of peptide synthesis before Section 3 gives an overview of state-of-the-art quality controls that are currently performed in each synthesis step. Section 4 then introduces our new quality control system from the design of the image acquisition setup over full automation to experiments that demonstrate the functionality. Subsequently, Section 5 deals with further applications as well as extensions and modifications of the quality control system that are planned for future research. The conclusion in Section 6 summarises the results gained and discusses the assets and drawbacks of the method.

2. Particle based solid phase peptide synthesis

Since the first peptide synthesis in 1882 by T. Curtius, peptides can be artificially produced by several methods, all basing on the principle of selectively concatenating amino acids of different kinds to create the desired peptides. The most common methods are liquid phase synthesis and solid phase synthesis. The former assumes that peptides are synthesised in solution (i.e. the growing peptides are free to move within the liquid), while the latter requires a solid support with fixed coupling sites to which the first amino acids can couple, so that the growing peptides are bound to their synthesis sites. By design, solid phase peptide synthesis surpasses the method based on liquids in the percentage of correctly assembled peptides as separation and purification of the desired peptide sequences are easier to perform.

Using a solid support, the medium that transports the amino acids to the growing peptides has to be determined. Selecting liquids as transport medium is a commonly used standard Fodor et al. (1991); Volkmer (2009). Nevertheless, it bears some disadvantages: First of all, the form and extension of a liquid drop on a support is hard to control. This means that synthesis spots for different peptides must be positioned on the support in a distance far enough that drops to build different peptides cannot overlap accidentally. Furthermore, there is a minimum size of the drops below which the drop evaporates before it can reach the surface and, at last, there is no control over the amino acids in the drop as they couple to the growing peptide upon contact to the support. This means that incorrectly deposited drops will immediately lead to incorrectly assembled peptides at that spot. In order to overcome these disadvantages, we developed a particle based transport Beyer et al. (2007) that embeds the amino acids in micro

particles. Thus, the coupling of the amino acids to the growing peptides is initiated after heating. Along with this new form of transport, the glass slide as solid support was replaced by a CMOS chip König et al. (2010) which facilitates this method of transport as described below.

Hence, our particle based solid phase peptide synthesis consists of the following four processing steps which are also depicted in Figure 1:

1. **Programming of the chip**
   Having equipped the CMOS chip with a chemical surface activation that provides the first coupling sites Stadler et al. (2007), the peptide synthesis can start. On the CMOS chip, each synthesis site is represented by an electrode that can be set up to a positive voltage of 100V. In the first step, all electrodes that are supposed to receive the same sort of amino acids are set to high voltage (usually 100V). All other electrodes are set to 0V, i.e. switched off.
2. **Particle deposition**
   Then, the particles that contain this kind of amino acids are deposited onto the chip: in a special deposition apparatus Wagner et al. (2011), the particles are negatively charged, so that they are guided to their synthesis sites by the electrical field generated by the electrodes, as also described in Löffler et al. (2011).
   After successful deposition of the first kind of amino acids, the chip is reprogrammed for the second kind and particle deposition is repeated using particles containing the second kind of amino acids.
3. **Coupling of amino acids**
   When each electrode is covered with particles, i.e. an entire layer in the peptide synthesis is completed, the chip is heated in an oven until the particles become gel-like. This initiates the coupling of the amino acids contained therein to the coupling sites on the surface activation of the chip.
4. **Chemical washing**
   To prepare the chip for the next layer of amino acids, chemical washing steps provide capping for those coupling sites that did not receive an amino acid, discard the remnants of the particles and activate the freshly deposited amino acids as coupling sites for subsequent depositions.
   Repeating the entire procedure, the amino acids are concatenated layer after layer, until the growing peptides reach their final length.

Fig. 1. Left to right: Solid phase peptide synthesis on CMOS chips: Negatively charged micro particles containing amino acids are guided to their synthesis sites by the selectively programmable electric fields of the CMOS chip, particles containing one specific sort of amino acids are deposited, particles containing all sorts of amino acids for the first layer of the peptide array are deposited, heating initiates the coupling of the first layer of amino acids to the surface, remnants of particles and excess amino acids are discarded in chemical washing, repeating this procedure leads to a complete peptide array (From Beyer et al. (2007). Reprinted with permission from AAAS.)

3. Related work on state-of-the-art quality control for peptide synthesis

Quality checks on each step of the peptide synthesis process are regularly performed. The size of the particles can be measured by means of a Malvern Mastersizer and their form by raster electron microscopy. The percentages of the particle contents are monitored by high-performance liquid chromatorgraphy (HPLC) to ensure that there is an excess of amino acids compared to the coupling sites available on the chip surface.

The deposition quality is checked by experts that investigate the spots on the support under a light microscope to find contaminations and to determine whether enough particles are deposited on each synthesis spot to guarantee a good coupling efficiency. If the spots are homogenously covered with particles, the experts assume that this deposition will result in a good coupling efficiency, having already achieved good synthesis results with such depositions, as demonstrated in Beyer et al. (2007); König et al. (2010).

The coupling efficiency itself can be estimated in the chemical washing step, when the freshly coupled amino acids are prepared to be the coupling sites for the next layer of amino acids. Then, protection groups (Fmoc groups) are removed from the amino acid that has just coupled to the growing peptide. The reaction product of the Fmoc group with the washing solution absorbs UV-light, so that UV spectrometry of the solution yields an estimate of coupled amino acids on the entire support. Yet, this gives no information about the number of coupled amino acids per synthesis spot, only averaged over the entire array. Furthermore, experiments with UV-spectrometry showed high error rates that originated from the fact that the measurements were at the detection limit of the spectrometer. Hence, at this degree of miniaturisation, the amount of reagents already becomes to small to be measured with standard chemical detection reactions.

After termination of the entire process, a selective antibody with a fluorescing molecule on top can be coupled to each peptide to measure the fluorescence signal as evidence for a correct synthesis. Furthermore, there is the possibility to perform a mass spectrometric analysis to investigate the synthesis quality of the assembled peptides. Both reactions are performed routinely Beyer et al. (2007) and are also standard for other types of peptide synthesis Roepstorff (2000). However, as the UV-spectrometry, these measurements are not spatially resolved, i.e. they only provide results for the entire support.

4. Quality control through image acquisition and processing

4.1 Motivation and outline

As discussed in Section 3, the standard quality analysis methods to determine the amount of coupled amino acids or the density of assembled peptides do not yield information about single synthesis spots but only for the entire array. In order to overcome this shortcoming, we developed a quality analysis setup that allows to evaluate single synthesis sites. The method is based on the correlation of two image acquisitions, namely, relating the image of a particle deposition with the respective image of coupled amino acids subsequently acquired by surface analysis, as described further below.

But before the correlation can be established, an image acquisition setup has to be assembled, as described in Section 4.2. Then, we set up a model for particle deposition using the image acquisition setup to compare the theoretical results with experimental tests. As discussed in

Section 4.3, this leads to improvements in the deposition step. Thus, having a tool for particle deposition analysis, Section 4.4 focuses on finding a suitable surface analysis technique that can image the chip surface retrieving signals from coupled amino acids with a resolution below the size of a synthesis spot (i.e. 100μm for the latest chip generation). Since the chosen surface analysis technique, time of flight secondary ion mass spectrometry (TOF-SIMS), destroys the surface during analysis, it cannot be included as quality analysis routine in the peptide array production process. Therefore, the particles deposited on the synthesis spot and the signals from the coupled amino acids in TOF-SIMS imaging have to be correlated to obtain a destruction free quality analysis. In Section 4.5, we define a quality measure based on this correlation that is capable of estimating the quality of single synthesis sites by examining the deposition image. Using this quality measure, we implemented an automated quality control, as described in Section 4.6, that reliably decides whether the density of coupled amino acids is sufficient to yield a predefined density of correctly assembled peptides or not and conveys this message to the peptide array production automaton that can react to the result.

4.2 Image acquisition setup for particle deposition

An image acquisition setup for particle deposition analysis should fulfil the following requirements in order to be suitable for quality analysis:

- The resolution of the entire setup (objective and camera) should be able to resolve the average diameter of the micro particles. Hence, the resolution should be in the range of 1-2μm Wagner et al. (2011).
- The optical imaging should be contactless, without interference of additional material, as for instance, oil immersion.
- If only a chip detail can be photographed, it should contain a satisfactory amount of synthesis sites to obtain a quality analysis evaluation that is representative for the entire chip. Since the experiments in this section are performed without an automated framework, chip details containing 100-160 synthesis sites, i.e. about 1% of the total amount of spots are photographed, as manual control samples showed that this amount already represents the entire chip well.
- The setup should be robust in handling so that all images are taken under comparable conditions, including a uniform illumination of the chip. To monitor the exact deposition for several amino acids, an alignment of the chip is also necessary. Yet, for retrieving a quality evaluation per synthesis site, aligning the images is not required.

Taking into account these requirements, we assembled an image acquisition setup consisting of a conventional inverted light microscope, a Zeiss Axiovert 35, a 10x/0.25 objective, a TV-adaptation of magnification 0.63, and a Progres C5 camera, as shown in Figure 2 on the left. The right hand side of Figure 2 shows a typical image taken from a particle deposition in which particles are depicted in white on the dark grey synthesis sites which are separated by a grid in brown.

4.3 Image processing supported deposition modelling

The image acquisition setup as assembled in Section 4.2 is used to take photos of particle depositions. Thereby, it was noticed that the deposition quality decisively depends on the programmed deposition pattern Löffler et al. (2011). Figure 3 shows some examples: the

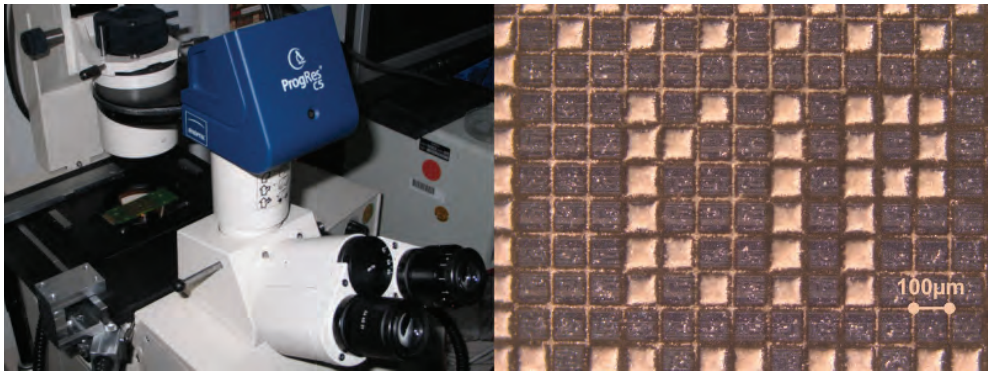


Fig. 2. Left: Image acquisition setup to monitor the particle deposition: Zeiss Axiovert 35, 10x/0.25 objective, TV-adaptation with magnification 0.63 and Progres C5 camera Right: Typical particle deposition image: particles (white), synthesis sites (dark grey), grid for separation of synthesis sites (brown)

pincushion effect on the top of the K, the central vertical line-like deposition in the I and the central horizontal line-like deposition in the P.

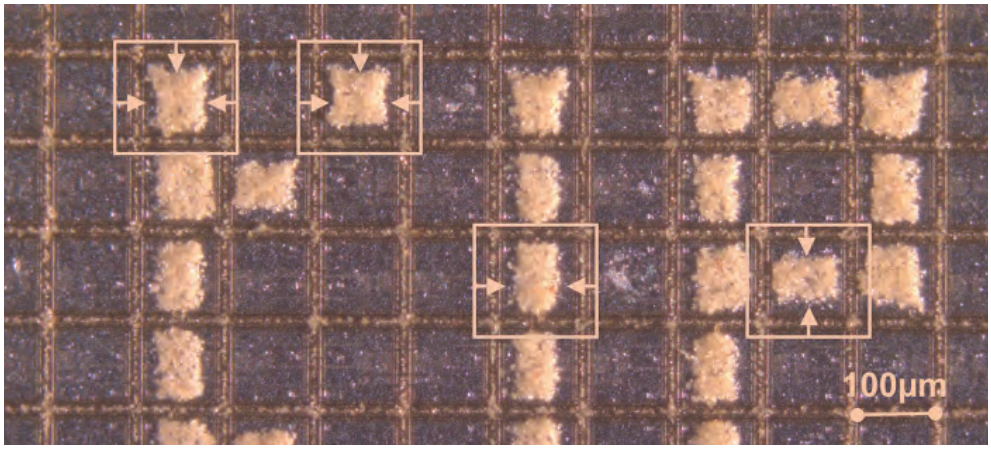


Fig. 3. The quality of the deposition decisively depends on the programmed pattern, the image shows the pincushion effect on the top of the K, the central vertical line-like deposition in the I and the central horizontal line-like deposition in the P

The effects could be explained by modelling the deposition process as a two-phase flow, with air as transport phase that carries the particles, subject to the boundary conditions in the deposition apparatus and the electrical field in front of the chip. Based on this model, particle deposition simulations were performed whose results correlated well with the effects observed in the images of actual depositions Wagner et al. (2011).

As a result, particle deposition patterns were modified such that undesired artefacts like the pincushion effect do not occur. This is achieved by splitting the deposition pattern into several

deposition steps for the same kind of amino acids, yielding high quality depositions as shown in Figure 2 on the right.

4.4 Spatially resolved image acquisition and processing of chemical surface analysis

As outlined in Section 3, chemical detection reactions are less precise and on the verge of the detection limit. This is not only the case for the Fmoc detection as described in Section 3, but also for the Bromophenol blue reaction, which is a standard reagent to visualise amino groups.

Hence, surface analysis techniques as listed in Table 1 are considered. The decisive parameters of the methods are the detection range, their analytical spot size and the profile depth. Requiring an analytical spot size in the range of 1 to 10μm, time of flight secondary ion

method	**detection range**	**lateral resolution**	**profile depth**
AES	10^{20} atoms/cm^3	1nm – 10nm	$<$ 10nm
AFM	1 atom/cm^3	100pm – 50pm	uppermost atom layer
TOF-SIMS	10^{15} atoms/cm^3	100nm – 50μm	$<$ 3nm
XPS	10^{19} atoms/cm^3	10μm – 1mm	$<$ 10nm

Table 1. Surface analysis techniques that may be able to resolve the density of amino acids on the chip surface: Auger electron spectroscopy (AES), atomic force microscopy (AFM), time-of-flight secondary ion mass spectrometry (TOF-SIMS), X-ray photoelectron spectroscopy (XPS)

mass spectrometry (TOF-SIMS) seems to be the best option. First of all it can detect down to 10^{15} atoms per cm^3, which should be sufficient to detect amino acids at the coupling sites on the surface because the surface density of coupling sites was determined to be in the range of 10^{16} per cm^2 Schirwitz et al. (2009). Secondly, its profile depth is small enough to analyse the uppermost layer of the surface only. Intruding deeper into the surface, artefacts from the chemical surface activation or even from the chip might obscure the signal actually coming from the amino acids at the top of the growing peptide. This excludes techniques as, for instance, X-ray photoelectron spectroscopy (XPS) or Auger electron spectroscopy (AES). Scanning probe microscopy methods (SPM), as atomic force microscopy (AFM), could also be considered due to their shallow profile depth and their capability of analysing spot sizes far less than 10μm. Yet, inherent to its functional principle, AFM cannot yield a good spatial resolution for the observation of single amino acids coupled to the surface, as the growing peptide is free to move vertically, leading to distorted images. Not shown in Table 1, spatially resolved matrix-assisted laser deposition/ ionisation (MALDI) could also be considered. Yet, this method is excluded again, as the minimum spot size analysable is 20μm.

The TOF-SIMS method is thus selected as method of choice, further encouraged by literature, demonstrating that single molecule detection, especially amino acid detection is possible with this method, for example in Aoyagi et al. (2008) or Hagenhoff (2000). Furthermore, it is possible to chemically characterise self-assembled monolayers (SAMs) by means of TOF-SIMS Graham & Ratner (1994), which can become important for peptide synthesis on CMOS chips as well, since SAMs are currently being established as chemically active surface coating. Knowing the TOF-SIMS fingerprint of the surface, it is easier to filter out the signals from the amino acids thereon.

The imaging of TOF-SIMS measurements works as depicted in Figure 4. The surface of the sample to be analysed is exposed to a pulsed beam of primary ions (usually gallium,

bismuth or noble gas ions). The secondary ions that are dissolved from the surface of the sample by the primary ions are accelerated according to their charge and enter a time-of-flight mass spectrometer in which they are separated according to their mass. This procedure is performed as a scan over the entire surface, so that a spatially resolved image of the detected mass distribution is created. The difficulty of recognising the amino acids among all other

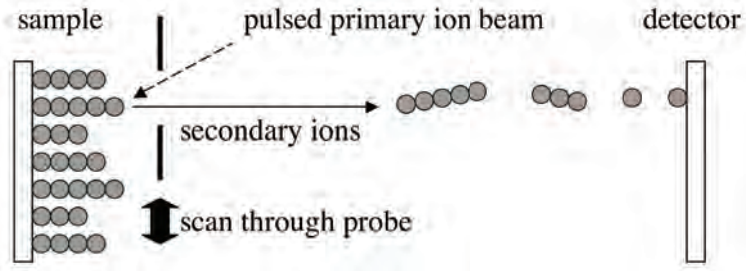


Fig. 4. TOF-SIMS: The sample is exposed to a pulsed beam of primary ions to dissolve secondary ions from the sample surface. The latter enter a time-of-flight mass spectrometer in which they are separated by their mass. Performing the procedure as a scan over the entire surface, a spatially resolved mass distribution is created

signals lies in the fact that, depending on the energy of the primary beam, the fragments ionised out of the sample surface can consist of any sequence of molecules that occurs in the sample.

For the TOF-SIMS imaging, an analysis chip with a test pattern containing three amino acids (proline, tryptophan and arginine) was created. The chip surface was chemically activated with coupling sites, the amino acids were deposited and the chemical washing steps performed, so that the imaging should yield signals from coupled amino acids on the surface. The measurements were then performed, using Bi_3^+ ions, producing images with a resolution of 128×128 pixels, i.e. one pixel corresponds to 3.91μm on a synthesis site.

Figure 5 shows the results: on the left hand side, the total ion rate indicates that all synthesis sites were homogeneously exposed to the ion beam, the centre image shows the most prominent scattered fragment from tryptophan (C_9H_8N), while the potassium signals on the right hand side clearly show the grid that separates the synthesis sites. The latter originates from the fact that the chemical surface activation is broken at the grid and thus confined to the synthesis spots, which was also an important observation discovered in the TOF-SIMS imaging. Hence, as the images show, the signal of the coupled amino acids is clearly visible despite the low signal to noise ratio of 2:1. Furthermore, signals from silicone oils were detected, originating from insufficiencies in the washing process when removing the particle remnants, so that this step could also be improved due to TOF-SIMS imaging results.

4.5 Destruction free quality control

4.5.1 Correlation between particle deposition and TOF-SIMS imaging

Having accomplished the imaging of coupled amino acids, it is possible to test the validity of the following experts' assumptions:

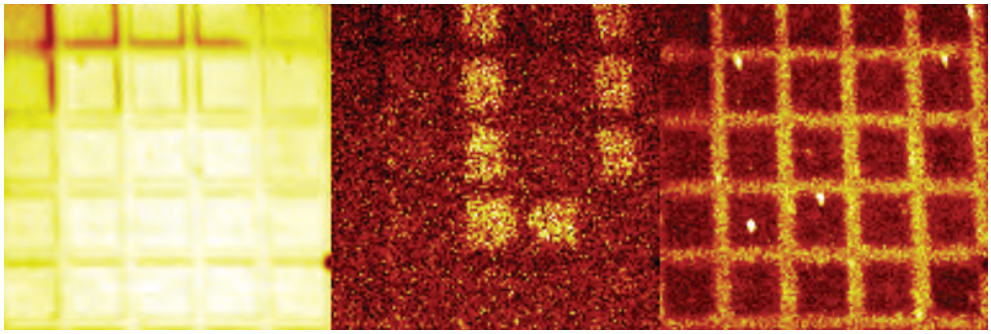


Fig. 5. TOF-SIMS images: Left: total ion image Centre: tryptophan signal with signal to noise ratio 2:1 Right: potassium signal indicating that the surface activation is confined to the spots

- A coverage with particles means that amino acids will couple to the surface underneath, i.e. a particle deposition on the synthesis sites automatically implies the coupling of amino acids there.
- During coupling, the melt of the particles flows isotropically into areas on the spot where no particles were deposited. The amino acids contained therein couple in those areas that the melt reaches.
- The particle depositions for each amino acid per layer must be performed twice to obtain the desired density of coupled amino acids.

To verify these assumptions, the correlation of the particle deposition images obtained with the image acquisition setup described in Section 4.2 with the TOF-SIMS images shown in Section 4.4 is performed by means of image processing.

First, both images are converted to greyscale. Then, the TOF-SIMS images are enlarged by Lanczos-2-kernel for optimal interpolation to match the size of the deposition images.

Subsequently, the two images are manually aligned and matched. Due to the low contrast in the grid regions of the tryptophan image, the image containing the potassium signal is used to match the grid in the TOF-SIMS images with the grid in the deposition images. Another option for alignment could have been to match the chip cell contents of the two images directly by applying a registration and matching algorithm, e.g. finding the largest overlap of the segmented foreground in both images. Yet, as the foreground is the object under investigation, it cannot be used for image alignment. Using the grid for alignment thus enables to detect a systematic shift in deposited particles when melting and coupling. This effect may be caused by unbalanced chip fixation on the circuit board or improper positioning in the oven during coupling, resulting in a unidirectional melt flow instead of a uniform distribution.

Next, both images are segmented by thresholding. The latter is calculated as the average of the maximum mean grey value for an uncovered synthesis site and the minimum mean grey value for a synthesis site covered with particles as this has proven to work well for the unimodal intensity distributions in the images Wagner et al. (2010).

The last step is the correlation of the segmented images, yielding that the first assumption is correct in principle. However, there are areas in which a particle deposition did not lead to a coupling of amino acids. This can be an artefact of the high noise or of the enlargement

of the TOF-SIMS image, but could be caused by deficiencies in the surface coating or steric hindrance to coupling as well. Furthermore, the second assumption is also true. The particle melt actually flows isotropically up to 5μm from the location of the deposition and induces coupling of amino acids in those areas. While the first two experts' assumptions were correct, the last one is fortunately not. As the images show contiguous signals from coupled amino acids on the surface, a second deposition step with the same sort of amino acids becomes obsolete, which saves deposition time and reduces the amount of required particles by one half.

4.5.2 Quality control system

Based on these results, a quality analysis was set up that uses the particle deposition image to determine whether the amount of amino acids expected to couple after this deposition is sufficient to continue with the next step or not. The algorithm to evaluate the quality of the synthesis sites on the deposition image was written in MATLAB and consists of the steps shown in Figure 6. Since the size of the synthesis sites has to be known in advance, each image acquisition setup requires a training step to be executed before quality analysis, in which the spot size is determined together with the resolution of the setup and a classification threshold to separate covered from uncovered spots, as well as the segmentation threshold that divides the particles from the background. The latter only has to be determined if the segmentation is performed by thresholding in the first place. Since these parameters only change when the setup is altered, the training step has to be performed once per setup configuration.

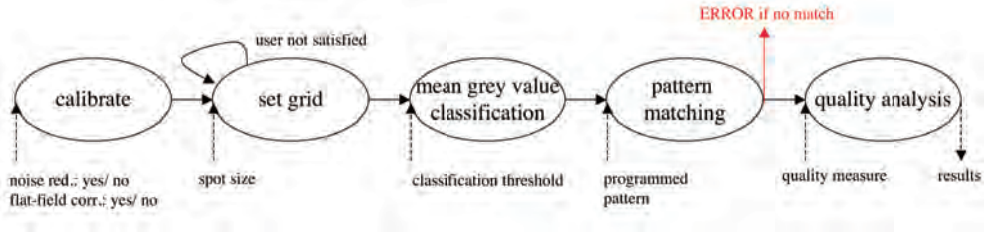


Fig. 6. Quality analysis algorithm: preprocessing is performed when necessary, set grid separates the image into the single synthesis spots, then a consistency check is performed by comparing the actual programming pattern with a rough classification of the spots before quality analysis of each spot starts

In the quality analysis algorithm, preprocessing like noise reduction or flat-field correction is performed when necessary. Subsequently, the image is separated into the single synthesis spots, setting the grid with one mouse click by the user. Later on, this can be performed by the synthesis automaton. Having extracted each single spot, a fast consistency check is performed, comparing the actual programming pattern with a rough classification of the spots. The classification calculates the mean grey value for each synthesis site and uses the trained threshold to classify the spots as uncovered or covered. Comparing the actual deposition pattern with this classification, grave errors like an incorrectly programmed electrode that require a restart in the deposition can be detected very quickly. If the consistency check terminates successfully, the actual quality analysis starts to evaluate each synthesis spot in the image.

As input, the quality analysis step requires a quality measure that decides whether the particle deposition is sufficient to continue the synthesis or not. This measure is set by user input, so that it meets the requirements set by the application the array is produced for. Knowing the correlation between particle deposition and coupling of the amino acids contained therein, the quality measure is defined by determining the degree of particle coverage of two central circles within the synthesis sites, as Table 2 summarises. The large outer circle covers 50% of the spot, while the smaller inner one covers 15%. A distinction is made between covered and uncovered spots because for the ones to be covered, the algorithm is supposed to determine whether the degree of coverage is sufficient to continue, while for the ones to be left uncovered, the amount of incorrectly deposited particles is decisive. Experimental tests using this quality measure

spot type	**good**	**satisfactory**	**bad**
covered	$>$ 90% of inner circle cov. $>$ 50% of outer circle cov.	$>$ 90% of inner circle cov. $>$ 0% of outer circle cov.	other
uncovered	$<$ 10% of inner circle cov. $<$ 20% of outer circle cov.	$<$ 20% of inner circle cov. $<$ 30% of outer circle cov.	other

Table 2. Quality measure based on two central circles for one spot Wagner et al. (2010)

showed that it works well for single particle depositions, as comparisons between human classification and automatic labelling showed high coincidences Wagner et al. (2010). Thus, the algorithm was prepared to cope with an entire peptide synthesis, analysing subsequent particle depositions per layer. This was done by verifying that two images of two subsequent particle depositions are consistent in so far, as the quality classification after the second deposition remains the same as after the first deposition for synthesis spots that were not affected by the second deposition.

Finally, the quality analysis algorithm was tested on a complete peptide synthesis, leading to classifications of 86.3% *good* spots, 5.0% *satisfactory* spots and 8.7% *bad* spots. In order to prove that the quality of the peptide array produced coincides with the prediction of the algorithm, a standard fluorescence detection on the peptides was performed. The image acquisition was performed with fluorescence scanner at a resolution of 5μm. Comparing the average classification label for each spot as shown in Figure 7 with the fluorescence intensity for the correctly assembled peptides, a high coincidence can be observed. The gradients in the left part of the fluorescence image, however, are not correctly predicted by the class labels. Yet, investigations over the entire array revealed that the insufficiencies in coupling at those sites were due to deficiencies in the chemical surface coating that were not subject to the quality control algorithm and hence not predictable.

The successful analysis of the peptide synthesis thus completes the quality analysis algorithm and arises the question how to actively control a peptide synthesis. Using the results shown in Figure 7, the threshold for the repetition of one deposition can be empirically trained, requiring more than 80% of all class labels to be *good* and not more than 20% of the class labels be *bad*. Furthermore, assigning more than 80% of all synthesis spots to the class *satisfactory* should also cause a warning because, in this case, the algorithm is not capable of finding a clear classification, since the *satisfactory* class is the one that is assigned in ambiguous cases.

Taking into account the subsequent particle depositions as well, the quality control algorithm should check the deposition quality of an entire layer by determining the ratio of *good* spots after the first deposition that remain *good* ones in the subsequent steps. The practical

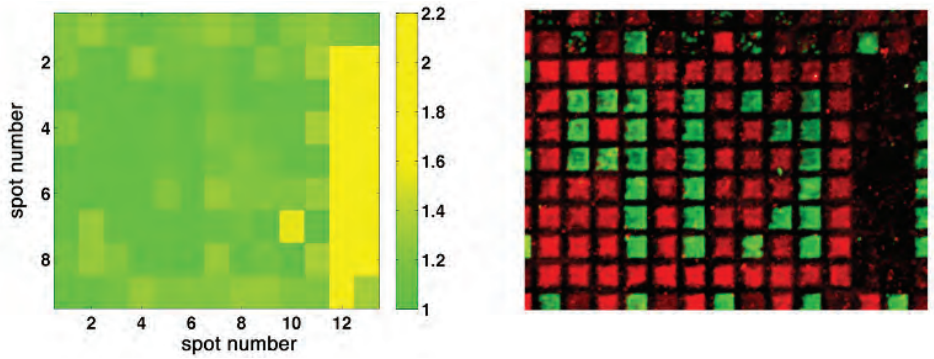


Fig. 7. Comparison of the average quality class with the fluorescence detection of the correctly assembled peptides for an entire peptide synthesis, class label 1 means *good*, class 2 label means *satisfactory* as defined in Table 2

experience of the analysed synthesis favours a threshold of 80% *good* $\rightarrow$ *good* label assignments for two subsequent depositions per layer, below which a repetition of the particle deposition for the entire layer becomes necessary.

4.6 Automation

With the quality control thresholds defined, the algorithm was implemented such that it automatically reads in images coming from the peptide synthesis automaton and outputs the control signals that operate the automaton according to the deposition quality. Since the analysis software is written in MATLAB while the machine control of the automaton still to be constructed is likely to be programmed in LabView, an interface has to be set up for the two programs to communicate with each other. Exchanging command and response files over a shared folder is the simplest method that assures compatibility even in the case when the quality analysis routine is translated to another programming language, e.g. to a C++ program. The LabView process control is supposed to convey the voltage programming files of the chip and the images taken by the camera to the quality analysis program, while the latter sends back responses about the quality to determine the next processing step. Furthermore, the LabView control program is supposed to store the parameter values how to position the chip above the image acquisition setup in order to take optimal pictures. Due to the high positioning accuracy, it is possible to fix the origin of the grid coordinates in the image, so that no user input is required to divide the image into the single synthesis spots before quality analysis.

The state machine shown in Figure 8 depicts the procedure of the automated quality control, i.e. started once, no user interaction is required during normal operation.

First, MATLAB and LabView are started. Before the automated particle deposition begins, LabView exports the chip programming deposition pattern into a folder shared by MATLAB. While LabView controls the particle deposition and takes the image for quality analysis, MATLAB scans that shared folder in previously defined time steps until LabView has exported the image to that folder. During the MATLAB analysis, LabView is in a wait state until it retrieves a response about the quality of the current deposition from MATLAB. The

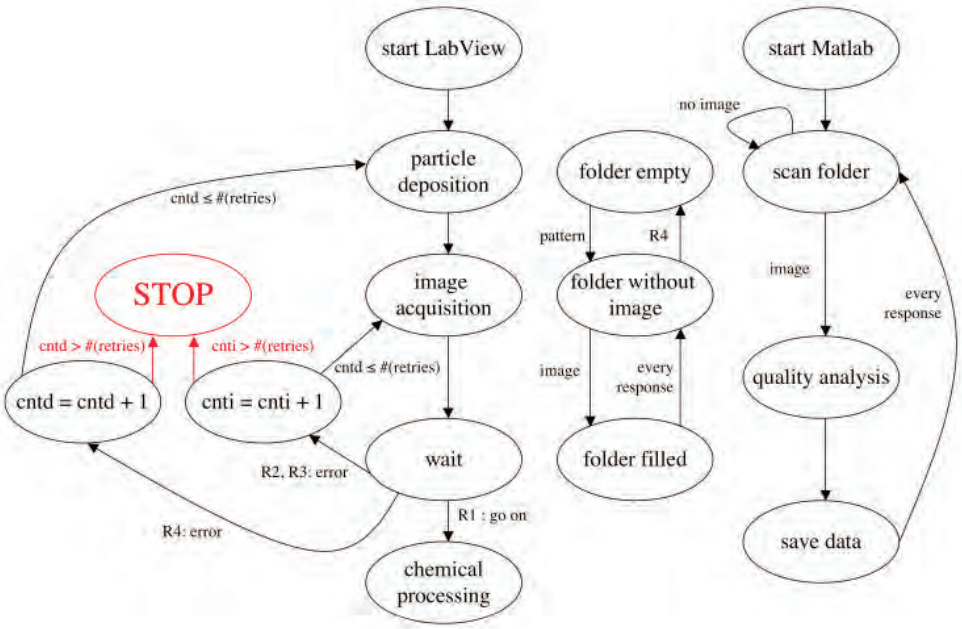


Fig. 8. Finite state machine diagram of automated quality control: after starting LabView, particle deposition, image acquisition, quality analysis and chemical processing can be performed and the respective actions in the image folder and in the MATLAB routine take place as described in the text

chip can already be brought back to the deposition unit of the automaton in the mean time, so that the next step can be immediately started after receiving the response – in the normal case, it is assumed that the quality will be sufficient to continue. The response, together with the analysed data, is saved into another shared folder from where LabView gets one of the four possible responses:

R1 : proceed to next step

R2 : the image data is corrupted so that it cannot be processed

R3 : the image causes a MATLAB intrinsic error

R4 : the image shows that the deposition is not of expected quality

In case of R1, LabView can go on with the next step in the processing queue. For the error messages R2 and R3, the problem could be caused in the image acquisition step. Therefore, this step is repeated once again with a new picture of the same particle deposition on the chip, while the deposition pattern file is kept. Receiving R4, the process has to clean the chip and repeat the steps from particle deposition onwards. In order to assure that no programming error of the chip caused the malfunction, the deposition pattern is also created anew and sent

to the shared folder. R4 could also be split into two errors, R4a and R4b, R4a dealing with sparse particle coverage and forcing an additional deposition step and R4b handling the case of too many incorrectly deposited particles requiring the chip to be cleaned prior to a new deposition. If the results do not improve after a previously determined number of retries, LabView stops the processing for human intervention and error inspection. Regardless of the response, MATLAB returns to scanning the shared folder waiting for a new image after termination of the analysis of the previous one.

5. Outlook, improvements and further applications

Integration of the image acquisition setup for quality control into the peptide synthesis automaton still has to be performed, but can easily be achieved replacing the microscope by a suitable objective and illumination device only, so that the optimum conditions defined in Section 4.2 are fulfilled.

The results obtained in Section 4.4 are the first of their kind and demonstrate that it is possible to detect the density of a single layer of coupled amino acids. Nevertheless, the proof of principle experiment should be repeated with an increased signal to noise ratio at a higher resolution so that the TOF-SIMS images can be better correlated to the deposition images that have a higher resolution. Furthermore, they should be performed for all kinds of amino acids used in peptide synthesis. This series of experiments could give insights in the coupling behaviour of the different amino acids and lead to semi-quantitative measurements of their coupling efficiencies.

Additional to that, the quality measure of the control algorithm could be totally based on the results gained in Section 4.4, simulating the coupling of amino acids, taking into account the (estimated areas and amounts of) coupled amino acids in the preceding layer and directly calculating the density of correctly assembled peptides out of this information. Although being more precise, this quality measure would certainly require more processing time for the exact simulation of the coupling and the calculation of the peptide density. Thus, since the time for analysis should be minimised, the current modus of quality analysis probably excels over this precise one, which remains to be checked.

Despite the good results obtained in Wagner et al. (2010) and Section 4.5, the predictive power of the algorithm can be further improved as currently, the assignment of the class labels is too pessimistic, leading to more repeated depositions than actually are necessary. With an increasing number of analysed synthesis, however, the quality measure can be refined, such that the number of unnecessary re-depositions is reduced.

It is also possible to extend the algorithm to evaluate fluorescence images like Figure 7 on the right. For this task, the single colour channels of the image can be evaluated after establishing a quality measure that bases on the intensity signal of the synthesis spots. Exchanging the classification component in the existing system by this one is simple due to its modular structure. The rest of the algorithm can be reused, since the structure of the synthesis sites will also remain the same. Knowing the (integrated) intensities per synthesis spot also enables biological data evaluation like finding the peptide that binds best to a certain test molecule or comparing different peptides with respect to their coupling probability.

Adapting the image acquisition and image processing system for glass slides as solid supports, the quality control can also be used in the peptide synthesis by means of a peptide

printer as described in Beyer et al. (2007), another method for automated peptide array production developed by us.

6. Conclusion: Comprehensive imaging based quality control for peptide array production

As shown in this chapter, image acquisition and image processing are vital parts of miniaturised peptide array production, when standard bio-chemical methods reach their detection limit and physical surface analysis techniques destroy the sample. Furthermore, imaging of the sample allows for spatially resolved results while standard quality analysis techniques only yield information about the entire sample.

Four different methods of image acquisition were used to accomplish an automated quality control system for peptide array production: electron microscopy, light microscopy, fluorescence microscopy and time of flight secondary ion mass spectrometry imaging (TOF-SIMS imaging). Figure 9 summarises the design of the quality control system and shows which results were obtained by the individual imaging methods.

To monitor the shape of the micro particles that contain the amino acids electron microscopy was used. It was found out that the majority of the particles has a shape that can be approximated by a sphere, which is an important fact because the size measurement of the particles outputs correct results only if the particles under analysis are approximately spherical. The programming of the chip and the particle deposition are checked by light microscopy. As a result, the comparison of several particle deposition images revealed that the quality of the deposition depends on the pattern programmed. Modelling the deposition and simulating the process, these artefacts could be explained and removed by splitting the deposition pattern into several steps for one kind of amino acid. Then, TOF-SIMS imaging was performed to visualise the density of a single layer of coupled amino acids with an unprecedented resolution of less than 4μm, yielding the first spatially resolved detection of coupled amino acids in such a miniaturised peptide array. As this method destroys the surface, a correlation between the particle deposition and the TOF-SIMS images was established via image processing to obtain a destruction free predictor for the amount of coupled amino acids when analysing the respective particle deposition. The demonstration that this predictor can be applied as quality control in each deposition step was performed by comparing its results with a fluorescence image of the correctly assembled peptides.

Algorithmically, the advantage of the implemented quality control program lies in its flexible structure so that it can analyse any of the images coming from light microscopy, TOF-SIMS imaging or fluorescence microscopy by setting a few parameters only. The training program additionally implemented simplifies the parameter tuning and allows to save standard configurations that can be quickly loaded into the quality control program.

The main future improvements on the system are planned in form of further TOF-SIMS imaging experiments to gain detailed information about the coupling efficiencies of all amino acids and increase the precision of the current correlation between particle deposition and amino acid coupling.

Performing all steps as described in this chapter, the quality control can, in principle, be applied to any kind of peptide array production. Adapting the algorithm to analyse glass slides instead of CMOS chips, and performing the imaging experiments with the arrays

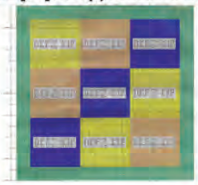
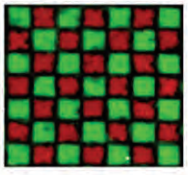
Particle creation	Image acquisition	Quality control
	*Electron microscopy* (monitoring the shape of the particles)	Routine checks guarantee reproducible experimental conditions (optimal particles are spheres)
Chip programming 	*Light microscopy* (consistency check for the correct deposition pattern)	Correct pattern, particle coverage and contamination can be monitored by light microscopy and image processing. The algorithm that decides to proceed or improve the deposition is based on the particle deposition simulations and the chemical analysis, correlating particle deposition with the density of coupled amino acids
Particle deposition 	*Light microscopy* (controlling the particle coverage and amount of incorrectly deposited particles)	
Chemical processing 	*TOF-SIMS imaging* (visualise density of coupled amino acids)	TOF-SIMS results lead to optimisation in chemical processing and development of the quality measure for the algorithm
Peptide array 	*Fluorescence microscopy* (visualising the density of correctly assembled peptides)	Reusing the algorithm for particle deposition, the number of correctly assembled peptides could be retrieved from the fluorescence image

Fig. 9. Summary of all peptide synthesis steps, the image acquisition methods used and the results obtained that lead to the setup of a quality control system

produced by a peptide printer Beyer et al. (2007), a quality control for another particle based solid phase peptide synthesis can be created. Using liquids as transport medium for the

amino acids, the system can still be applied, however, solely as quality analysis because coupling of the amino acids happens immediately when the liquid contacts the surface. Nevertheless, detecting deficiencies at an early synthesis stage saves production time and costs, as non-functional arrays can be discarded just after the step the deficiency occurs in and not at the end of an entire synthesis process.

7. Acknowledgements

We thank Ralf Achenbach, Markus Dorn and the ASIC laboratory for their technical assistance with the microchips, Daniela Rambow, Sebastian Heß, Jürgen Kretschmer, Thomas Felgenhauer and Volker Stadler for their assistance, help and advice concerning particle production and surface activation, Klaus Leibe for his contributions to the design of the deposition apparatus and Jürgen Hesser for stimulating discussions. We gratefully acknowledge the funding of the Baden-Württemberg-Stiftung and the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences.

8. References

Angulo, J. & Serra, J. (2003). Automatic analysis of DNA microarray images using mathematical morphology, *Bioinformatics* 19(5): 553–562.

Aoyagi, S., Rouleau, A. & Boireau, W. (2008). TOF-SIMS structural charaterization of self-assembly monolayer of cytochrome b5 onto gold substrate, *Applied Surface Science* 255: 1071–1074.

Bajcsy, P. (2005). An overview of DNA microarray image requirements for automated processing, *Computer Vision and Pattern Recognition Workshop* 0: 147.

Beyer, M., Nesterov, A., Block, I., König, K., Felgenhauer, T., Fernandez, S., Leibe, K., Torralba, G., Hausmann, M., Trunk, U., Lindenstruth, V., Bischoff, F. R., Stadler, V. & Breitling, F. (2007). Combinatorial synthesis of peptide arrays onto a microchip, *Science* 381(5858): 1888.

Brown, C. S., Goodwin, P. C. & Sorger, P. K. (2001). Image metrics in the statistical analysis of DNA microarray data, *Proceedings of the National Academy of Sciences* 98(16): 8944–8949.

Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X. C., Stern, D., Winkler, J., Lockhart, D. J., Morris, M. S. & Fodor, S. P. A. (1996). Accessing genetic information with high-density DNA arrays, *Science* 274(5287): 610–614.

Cretich, M., Damin, F., Pirri, G. & Chiari, M. (2006). Protein and peptide arrays: Recent trends and new directions, *Biomolecular Engineering* 23(2-3): 77–88.

Debouck, C. & Goodfellow, P. N. (1999). DNA microarrays in drug discovery and development., *Nature Genetics* 21(1 Suppl): 48–50.

Fodor, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. & Solas, D. (1991). Light-directed, spatially addressable parallel chemical synthesis, *Science* 251(4995): 767–773.

Graham, D. J. & Ratner, B. D. (1994). Multivariate analysis of TOF-SIMS spectra from dodecanethiol SAM assembly on gold: Spectral interpretation and TOF-SIMS fragmentation processes, *Science* 264(5157): 399–402.

Gusev, A. I., Wilkinson, W. R., Proctor, A. & Hercules, D. M. (1993). Quantitative analysis of peptides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Applied Spectroscopy* 47(8): 1091–1092.

Hagenhoff, B. (2000). High resolution surface analysis by TOF-SIMS, *Miicrochimica Acta* 132(2–4): 259–271.

Heller, M. J. (2002). DNA MICROARRAY TECHNOLOGY: Devices, systems, and applications, *Annual Review of Biomedical Engineering* 4(1): 129–153.

König, K., Block, I., Nesterov, A., Torralba, G., Fernandez, S., Felgenhauer, T., Leibe, K., Schirwitz, C., Löffler, F., Painke, F., Wagner, J., Trunk, U., Hausmann, M., Bischoff, F., Breitling, F., Stadler, V. & Lindenstruth, V. (2010). Programmable high-voltage cmos chips for particle-based high-density combinatorial peptide synthesis, *Sensors and Actuators B* 147(418): 418–427.

Löffler, F., Wagner, J., König, K., Märkle, F., Fernandez, S., Schirwitz, C., Torralba, G., Hausmann, M., Lindenstruth, V., Bischoff, F. R., Breitling, F. & Nesterov, A. (2011). High-precision combinatorial deposition of micro particle patterns on a microelectronic chip, *Aerosol Science and Technology* 45: 65–74.

Roepstorff, P. (2000). Maldi-tof mass spectrometry in protein chemistry, *EXS* 88.

Schirwitz, C., Block, I., König, K., Nesterov, A., Fernandez, S., Felgenhauer, T., Leibe, K., Torralba, G., Hausmann, M., Lindenstruth, V. & Stadler, V. (2009). Combinatorial peptide synthesis on a microchip, *Current Protocols in Protein Science* 57: 18.2.1–18.2.13.

Stadler, V., Beyer, M., König, K., Nesterov, A., Torralba, G., Lindenstruth, V., Hausmann, M., Bischoff, F. R. & Breitling, F. (2007). Multifunctional cmos microchip coatings for protein and peptide arrays, *Journal of Proteome Research* 6: 3197–3202.

Templin, M. F., Stoll, D., Schwenk, J. M., Pötz, O., Kramer, S. & Joos, T. O. (2003). Protein microarrays: Promising tools for proteomic research, *PROTEOMICS* 3(11): 2155–2166.

Volkmer, R. (2009). Synthesis and application of peptide arrays: Quo vadis SPOT technology, *ChemBioChem* 10(9): 1431–1442.

Wagner, J., König, K., Förtsch, T., Löffler, F., Fernandez, S., Felgenhauer, T., Painke, F., Torralba, G., Lindenstruth, V., Stadler, V., Bischoff, F., Breitling, F., Hausmann, M. & Nesterov-Müller, A. (2011). Microparticle transfer onto pixel electrodes of 45μm pitch on HV-CMOS chips – simulation and experiment, *Sensors and Actuators A. Physical* 172: 533–545.

Wagner, J., Löffler, F., König, K., Fernandez, S., Nesterov-Müller, A., Breitling, F., Bischoff, F., Stadler, V., Hausmann, M. & Lindenstruth, V. (2010). Quality analysis of selective microparticle deposition on electrically programmable surfaces, *Rev. Sci. Instrum.* 81(7): 073703–1–073703–6.

Wang, X., Ghosh, S. & Guo, S.-W. (2001). Quantitative quality control in microarray image processing and data acquisition, *Nucleic Acids Research* 29(15): e75.