Analyst

RSC Publishing

COMMUNICATION

View Article Online
View Journal | View Issue

Biocatalytic analysis of biomarkers for forensic identification of ethnicity between Caucasian and African American groups†

Cite this: *Analyst*, 2013, **138**, 6251

Received 27th May 2013
Accepted 12th August 2013

DOI: 10.1039/c3an01062g

www.rsc.org/analyst

Friederike Kramer,[ac] Lenka Halámková,[abe] Arshak Poghossian,[cd] Michael J. Schöning,[cd] Evgeny Katz*[a] and Jan Halámek*[ae]

A new biocatalytic assay analyzing the simultaneous presence of creatine kinase (CK) and lactate dehydrogenase (LDH) was developed aiming at the recognition of biofluids of different ethnic origins for forensic applications. Knowing the difference in the concentrations of CK and LDH in the blood of healthy adults of two ethnical groups, Caucasian (CA) and African American (AA), and taking into account the distribution pattern, we mimicked the samples of different ethnic origins with various CK–LDH concentrations. The analysis was performed using a multi-enzyme/multi-step biocatalytic cascade where the differences in both included enzymes resulted in an amplified difference in the final analytical response. The statistically established analytical results confirmed excellent probability to distinguish samples of different ethnic origins (CA *vs.* AA). The standard enzymatic assay routinely used in hospitals for the analysis of CK, performed for comparison, was not able to distinguish the difference in samples mimicking blood of different ethnic origins. The robustness of the proposed assay was successfully tested on dried/aged serum samples (up to 24 h) – in order to mimic real forensic situations. The results obtained on the model solutions were confirmed by the analysis of real serum samples collected from human subjects of different ethnic origins.

Forensic chemical/biochemical analysis is an essential tool in the criminal justice system, particularly when examining physical evidence to support criminal investigation and subsequent prosecution.[1] Forensic evidence that is carefully gathered and analyzed can be an important step leading to the arrest and conviction of a suspect. A biochemistry/molecular biology-based subarea of forensic analysis, forensic serology, deals with the complex task of gathering information on the type, age, origin or gender from biological fluids (blood, saliva, *etc.*) found on a crime scene.[2–6] If blood or other fluids not belonging to a victim are found at a crime scene, they can be analyzed giving important information for investigation. Such data can significantly improve the information pool about possible suspects. Modern forensic serology[7–9] relies on two major methods: immunoassay[9] and DNA[10]/RNA[11] analysis. Apart from very traditional, rather primitive, immunoprecipitation techniques,[12] which have been carried out for dozens of years and are used routinely for blood group determination ('blood typing'),[7,8] DNA analysis, a sophisticated approach based on polymerase chain reaction (PCR), electrophoresis, and blotting techniques, provides excellent, valuable, and complex results in the field of forensic analysis.[10] Nowadays, DNA-based techniques need to be employed for gathering complex sets of information, like gender, age or race from body fluid samples exclusively.[13] One such technique, 'DNA profiling', is used to obtain a DNA 'fingerprint' from a biological sample and compare it to profiles obtained from DNA at a crime scene, from an individual, or from profiles stored in a database.[14] Recent development in this area allowed for the faster apprehension of suspects through comparing newly obtained crime scene samples to those already stored in the database, thus providing links between specific criminals and other crimes. A second application of DNA analysis involves when a suspect is found. 'DNA matching' can often be used to either prosecute or release a person, as the blood and other bodily fluids can create a direct link between a violent crime scene and an assailant. However, most of these analyses are based on sophisticated techniques and require complex instrumentation.[7] Thus, current on-site analysis in forensic serology is composed entirely of simple identification of possible samples (*e.g.* bloodstains)[15] and the

[a]*Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, NY 13699, USA. E-mail: ekatz@clarkson.edu; jhalamek@clarkson.edu; Fax: +1-315-268-6610; Tel: +1-315-268-4421*

[b]*Department of Biology, Clarkson University, Potsdam, NY 13699, USA*

[c]*Institute of Nano and Biotechnologies, Aachen University of Applied Sciences, Campus Jülich, Heinrich-Mußmann-Str. 1, D-52428 Jülich, Germany*

[d]*Peter Grünberg Institute (PGI-8), Research Centre Jülich GmbH, D-52425 Jülich, Germany*

[e]*New affiliation from September 1st 2013: Department of Chemistry, University of Albany, State University of New York, 1400 Washington Avenue, Albany, NY, 12222. E-mail: jhalamek@albany.edu*

† Electronic supplementary information (ESI) available: Detailed Experimental section. See DOI: 10.1039/c3an01062g

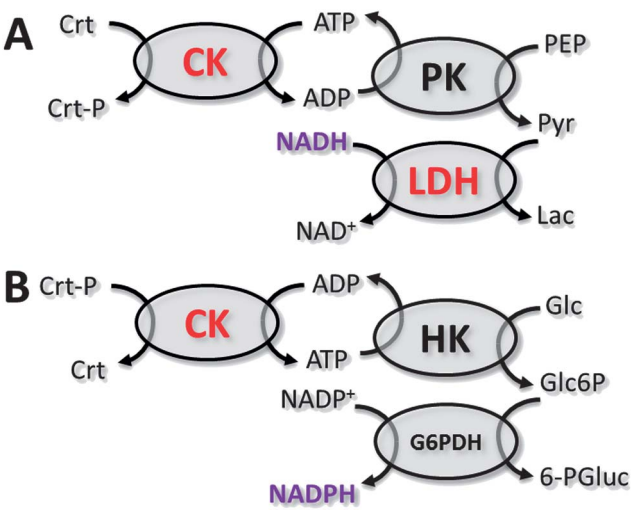
collecting steps, followed by transportation of these samples to specialized laboratories. This introduces time-delay in the investigation procedure and increases the complexity of the investigation itself. Despite the modern trend of designing portable equipment for on-site forensic biochemical analysis,[7,16] including microfluidic[17] and lab-on-a-chip[18] systems, there is an obvious lack of on-site detection/characterization technologies, analogous to point-of-care diagnostic approaches, which have become common in medical areas (*e.g.* diabetes management, pregnancy tests, *etc.*).

Recent advances in chemical[19] and biochemical[20] unconventional computing, particularly based on enzyme-catalyzed reactions,[21] allowed the formulation of biocatalytic cascades activated by biomarkers characteristic of various pathophysiological conditions (*e.g.* different injuries).[22–24] In general, this approach resulted in novel bioanalytical methods where combinations of biomolecular inputs resulted in simple diagnostic conclusions in the binary YES/NO format.[25,26] This approach, first tested in model systems,[22–24] has been successfully extended to animal studies[27] allowing the combination of several species of interest in a single analytical test. The developed analytical systems do not provide quantitative information about all analyzed species but they do allow a simple qualitative discrimination between two categories of samples (*e.g.* "healthy"/"unhealthy"). In the present research we extended this approach to forensic needs by analysis of biomarkers characteristic of different ethnic groups. It should be noted that many different biomolecular markers (particularly proteins/enzymes) vary in biofluids depending on the ethnic origin.[28,29] However, the analysis of most of them requires sophisticated proteomic methods and complex instrumentation (*e.g.* mass-spectroscopy).[30] Still, some of the enzymes can be analyzed using relatively simple assay procedures. In order to amplify the difference in their analytical discrimination the assay should include two or more analyzed biocatalytic species in the same multi-step biocatalytic cascade. Therefore, the biomarker selection should be based on two criteria: (i) significant difference in the concentrations depending on the sample origin (in our case depending on the ethnic origin) and (ii) the possibility to include them in a single biocatalytic cascade for amplifying their effect on the final analytical output signal.

Two enzymes, creatine kinase (CK) and lactate dehydrogenase (LDH), were selected for the analysis based on their known different concentration levels in the blood of healthy adults of two ethnic groups: Caucasian (CA) and African American (AA).[31,32] Mean concentrations of CK are 180 U L^{-1} and 665 U L^{-1} in CA and AA ethnic groups,[31] respectively (note a large difference of 485 U L^{-1}), while mean concentrations of LDH differ much less: 152 U L^{-1} and 167 U L^{-1} in CA and AA groups,[32] respectively (the difference of 15 U L^{-1} only). It is well known that CK[33–35] and LDH[36,37] serum levels can be used as diagnostic tools for various injuries[22–24] and their concentrations significantly vary for "healthy" and "unhealthy" samples, however, they have never been applied to the analysis of "healthy" samples with different ethnic origins. One of the reasons for this is because of a relatively small difference in their concentrations depending on the origin, while the person-to-person variation can easily screen the original difference. In order to solve this problem both biomarkers (CK and LDH) were included in a single multi-step biocatalytic cascade resulting in amplification of the difference. For illustrating the advantage of this approach, a single CK enzyme-assay[38–40] was applied for comparison.

Scheme 1A and B show the biocatalytic cascades used in the new two-enzyme CK–LDH-assay and single-enzyme CK-assay, respectively (see Experimental details in the ESI†). The CK–LDH-assay was experimentally optimized for the best performance, while the CK-assay was used in its standard version recommended by Sigma-Aldrich[41] and commonly used in hospitals.[42] For both analytical procedures we used sets of samples mimicking CK and LDH concentrations in CA and AA groups and then applied statistical analysis to demonstrate and evaluate the difference between the samples mimicking different ethnic (CA and AA) origins.

The CK concentration distribution in human plasma has been previously studied[31] in order to investigate the CK variability in baseline serum and the contribution of ethnicity, gender and other factors. The study has reported significant ethnic differences in CK levels. The reported data were used in the present study to prepare solutions mimicking CK levels in the blood of different ethnic origins (CA and AA). We started from the statistical analysis of the available data. The values were not normally distributed, but rather highly positively skewed and consistent with a log-normal distribution. The parameters of the log-normal distribution were available only for overall CK values, while the distribution parameters estimated from the CA and AA data groups came from logarithmic



Scheme 1 (A) The biocatalytic cascade for the two-enzyme CK–LDH-assay. (B) The biocatalytic cascade for the one-enzyme CK-assay. The following abbreviations are used in the scheme: CK (creatine kinase), PK (pyruvate kinase), LDH (lactate dehydrogenase), Crt (creatine), Crt-P (creatine phosphate), ATP (adenosine 5′-triphosphate), ADP (adenosine 5′-diphosphate), NAD^+ (β-nicotinamide adenine dinucleotide), NADH (β-nicotinamide adenine dinucleotide reduced), PEP (phospho(enol)pyruvic acid), Pyr (pyruvate), Lac (lactate), HK (hexokinase), G6PDH (glucose-6-phosphate dehydrogenase), $NADP^+$ (β-nicotinamide adenine dinucleotide phosphate), NADPH (β-nicotinamide adenine dinucleotide phosphate reduced), Glc (glucose), Glc6P (glucose-6-phosphate) and 6-PGluc (6-phosphate gluconic acid). The exact information on the materials is given in the ESI.†

untransformed data. To generate CK concentrations/values for CA and AA groups that arise from the same distribution reported in the CK study, we first corrected the only available parameters for normal distribution to log-normal. Using the standard R-project software,[43,44] we generated random values according to the recalculated parameters for the log-normal distribution of CA ($M_{log} = 4.37 \pm 1.28$ mU mL^{-1}) and AA ($M_{log} = 5.61 \pm 1.34$ mU mL^{-1}) groups based on the CK distribution parameters reported by Deuster *et al.*[31] We employed 25 randomly calculated concentrations for each group. The two sets with identical CK concentrations mimicking the CK distribution in CA and AA groups were then used in the two-biomarker CK–LDH-assay (see Scheme 1A) and then in the single-biomarker CK-assay (see Scheme 1B), for comparison. Hence, the two-biomarker enzymatic assay requires variable LDH input, another set of values (LDH concentrations) characteristic of CA ($M_{log} = 4.99 \pm 0.21$ mU mL^{-1}) and AA ($M_{log} = 5.09 \pm 0.24$ mU mL^{-1}) groups has been generated to follow a log normal distribution. Since the parameters have been reported for a normal distribution[32] they have been recalculated first to follow a log-normal distribution. Since these two biomarkers (CK and LDH) are not associated,[45,46] both sets of CK and LDH values have been randomly paired together by R-project software and these CK–LDH concentration pairings have been used in the analysis of the samples mimicking the distribution of CA and LDH concentrations for the CA and AA groups (see details in the ESI†).

Fig. 1 and 2 show the experimental results using the two-enzyme CK–LDH-assay and the single-enzyme CK-assay, respectively, where the CK and LDH concentrations were selected from the values known for the CA and AA ethnic groups with distribution parameters calculated according to the procedure described above. Note that the analytical responses are represented by the decreasing optical absorbance of NADH in the CK–LDH-assay and increasing optical absorbance of NADPH in the CK-assay, respectively, being consistent with the biocatalytic cascades used in the assays (see Scheme 1). To examine the distribution of our output data we drew histograms, Fig. 3 and 4. Since our input data exhibit a log-normal distribution, this distribution is projected into our output data to some extent and extreme values influence the shape of the distributions in all four plots. Additionally, we superimposed a probability density function (PDF) to our histograms.[47] Since all the distributions are skewed we favored the non-parametric approach over choosing an underlying distribution. The kernel density estimation, which is the most common non-parametric method,[48,49] has been applied (see details in the ESI†).

The presented histograms with the superimposed PDFs show how the output signals for CA and AA groups are distributed, *i.e.* how their separation will enable us to distinguish between these two groups. It is apparent from the illustrated histograms that the two-enzyme (CK–LDH) method performs much better than the standard single-enzyme (CK) method in terms of the overlap between two groups. The two-enzyme CK–LDH-assay achieves better results than the standard single-enzyme CK-assay also in terms of the distance between the distributions.[50,51] In other words, the separation of the output optical signals was significantly increased when two enzymes were analyzed in the biocatalytic cascade. When the output signal distributions have wide separation, the analytical discrimination between the analyzed CA and AA groups

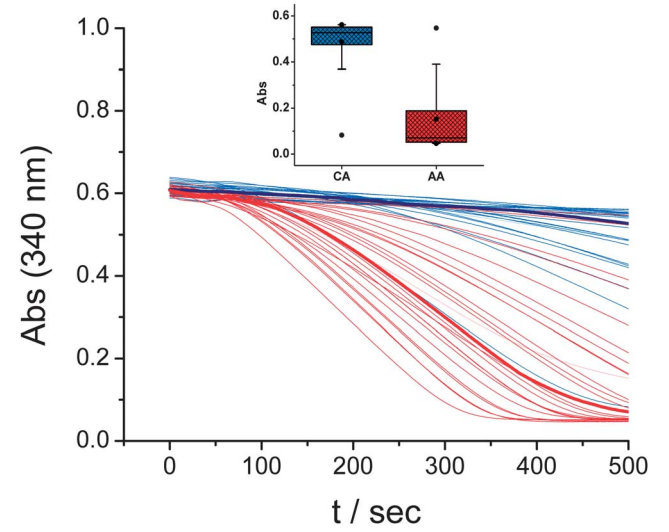


Fig. 1 Absorbance ($\lambda = 340$ nm) corresponding to the consumption of NADH upon operation of the CK–LDH-assay. The bottom (red) and top (blue) traces correspond to the application of samples with CK and LDH concentrations mimicking AA and CA groups, respectively. The bold solid curves show the median responses for both groups. Inset: box and whisker plot of Abs in AA and CA groups. The median value for each group is noted with the horizontal line in a box, the boxes represent the range of values from the 25th percentile to the 75th percentile, the ends of the whiskers represent the 5th and the 95th percentile of values, and the dots are the mean, maximum and minimum values.

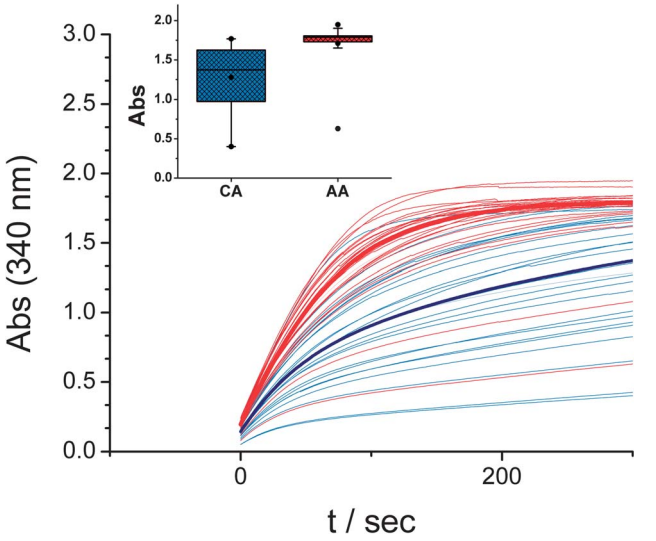


Fig. 2 Absorbance ($\lambda = 340$ nm) corresponding to the production of NADPH upon operation of the CK-assay. The bottom (blue) and top (red) traces correspond to the application of samples with CK concentrations mimicking CA and AA groups, respectively. The bold solid curves show the median responses for both groups. Inset: box and whisker plot of Abs in AA and CA groups. The median value for each group is noted with the horizontal line in a box, the boxes represent the range of values from the 25th percentile to the 75th percentile, the ends of the whiskers represent the 5th and the 95th percentile of values, and the dots are the mean, maximum and minimum values.

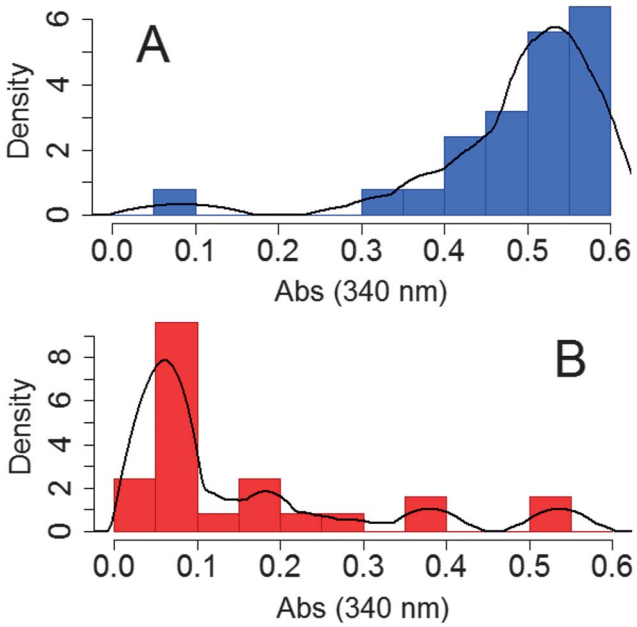


Fig. 3 Density histograms of the output signal (absorbance) obtained for: (A) CA group (blue color) and (B) AA group (red color) using the CK–LDH-assay (see the biocatalytic cascade shown in Scheme 1A). The histograms were derived from the experimental data shown in Fig. 1. Superimposed is the kernel density curve (black line).

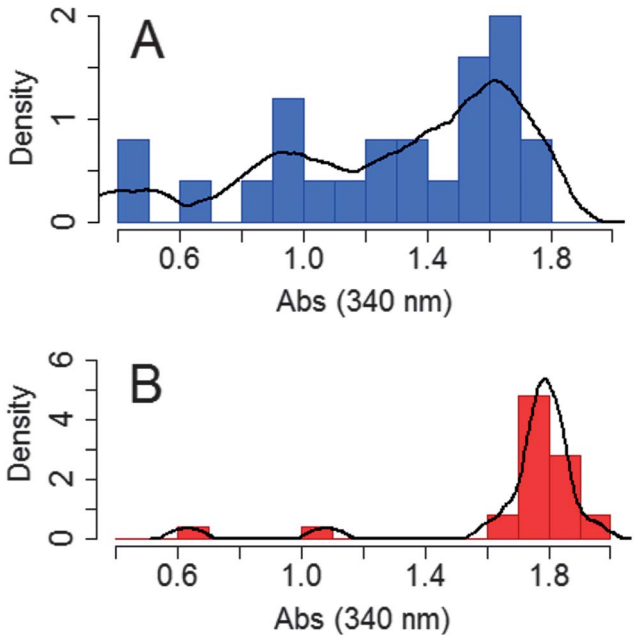


Fig. 4 Density histograms of the output signal (absorbance) obtained for: (A) CA group (blue color) and (B) AA group (red color) using the CK-assay (see the biocatalytic cascade shown in Scheme 1B). The histograms were derived from the experimental data shown in Fig. 2. Superimposed is the kernel density curve (black line).

becomes easier. In order to mathematically define the quantitative degree of separability of the two probability distributions over the full set of values metric dissimilarity/similarity measures have been calculated (see details in the ESI†). The Hellinger distance (HD) was used to quantify the dissimilarity between two probability distributions and the Bhattacharyya coefficient (BC) was applied as a measure of the similarity between two probability density functions.[50] A higher value of the BC is associated with more similarity between the histograms. The opposite statement is valid for HD, where a higher value is associated with more dissimilarity between the histograms. For the new two-enzyme CK–LDH-assay the BC and HD values were 0.34 and 0.81, respectively, while the standard single-enzyme CK-assay results in BC and HD equal to 0.83 and 0.41, respectively. An improvement in robustness of the new two-enzyme CK–LDH analytical assay has been demonstrated by its better applicability for distinguishing the difference between CA and AA analyzed groups.

The experiments described above confirmed the applicability of the two-enzyme CK–LDH-assay for distinguishing the difference between the CA and AA groups. They were performed in model solutions, however, comprised of an aqueous buffer spiked with the CK and LDH to the concentrations characteristic of the analyzed groups. Under real conditions of interest for forensic analysis, the biomarkers should be analyzed in bloodstains found at a crime scene. In order to mimic forensic conditions we performed the analysis in human serum solutions. In the first set of measurements we prepared the samples from serum representing mixed solutions from various donors commercially available from Sigma-Aldrich. Since the samples were mixtures from a number of different donors with averaged concentrations of CK and LDH biomarkers, we used the serum samples as they were to mimic the CA group, while we dissolved additional amounts of the CK (485 mU mL^{-1}) and LDH (15 mU mL^{-1}) biomarkers for mimicking the AA group to obtain the concentration difference typical for the AA group *vs.* the CA group. It should be noted that in this set of data the absolute values of the CK and LDH biomarkers might be different from the natural concentrations, but their difference was similar to that between the CA and AA groups. After preparing two sets of samples mimicking the CA and AA groups, the serum samples were dried and preserved at 35 °C for different time intervals (up to 24 hours), after which the samples were re-dissolved in an aqueous buffer containing the required enzyme substrates and then analyzed according to the two-enzyme CK–LDH-assay (see Scheme 1A; the experimental details are given in the ESI†). The results are shown in Fig. 5, where the zero-time interval corresponds to the analysis of freshly prepared samples without drying and all other time intervals correspond to different ageing of the dry serum stains prior to their analysis. Absorbance changes measured in the assay (similar to those shown in Fig. 1 for the model solutions) were normalized to the maximum value characteristic of the fresh samples mimicking the AA group. The assay demonstrated almost unchanged optical responses (meaning the same CK–LDH activity) in the samples mimicking the CA group over the ageing time intervals up to 24 hours. On the other hand, the samples mimicking the AA group demonstrated somewhat reduced optical changes (attributed to the expected decreasing CK–LDH activity) in the time intervals up to *ca.* 10 hours, while after that the optical responses reached a limit and stayed unchanged until the

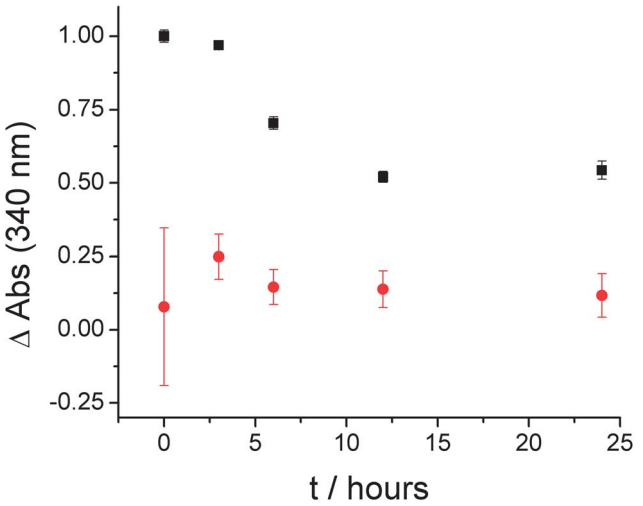


Fig. 5 Absorbance changes (ΔAbs) obtained for the two-enzyme CK–LDH-assay applied to the re-dissolved serum samples mimicking CA (red circles) and AA (black squares) groups after drying and ageing for different time intervals. The zero-time interval corresponds to the analysis of the freshly prepared samples without drying. The data represent mean values of ΔAbs normalized to the maximum ΔAbs value characteristic of the fresh samples mimicking the AA group; and the error bars represent relative standard errors of ΔAbs measurements from 5 samples. Note that the samples were composed of mixed serum from different donors with added CK and LDH to mimic their concentration difference in the CA and AA groups.

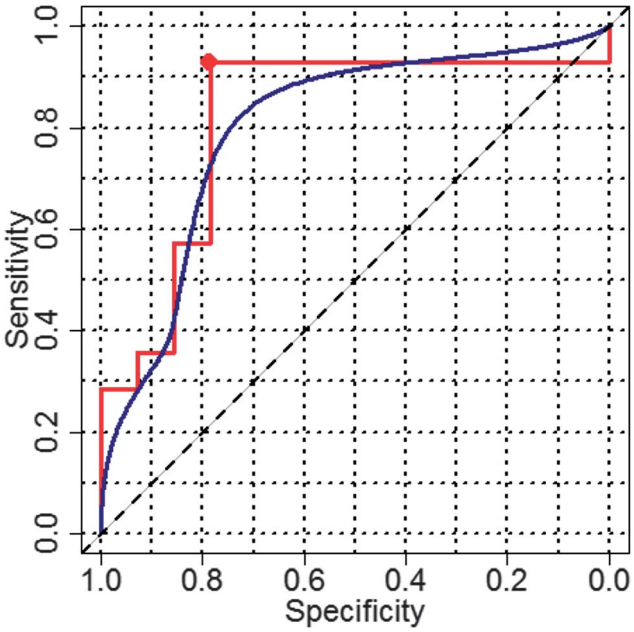


Fig. 6 Receiver operating characteristic (ROC) empirical (red) and smoothed (blue) curves for the two-enzyme CK–LDH-assay. Random choice is denoted by the diagonal line. The red point on the plot corresponds to the best sensitivity–specificity pair (the best tradeoff between them) giving the most accurate cut-off point for discrimination between CA and AA serum samples. Note that samples were serum from individual donors with known ethnic origins.

maximum ageing time was applied (24 hours). Despite the fact that the gap separating the analyzed CA and AA samples was decreased in 10 hours of the sample ageing by *ca.* 50%, the analytical responses from the CA and AA groups were perfectly distinguishable, thus confirming the applicability of the proposed analytical method for dry sample stains which are at least 24 hours old. In the present preliminary study we did not attempt ageing longer than 24 h for the dry samples. It should also be noted that in the preliminary step we analyzed only the samples with the mean concentration difference without taking into account the natural variability which can potentially screen the difference between the CA and AA groups.

In the next phase of the study, we analyzed serum samples obtained from individual donors with known ethnic origins (CA and AA). The samples were obtained from ProMedDx Specimen Bank (Norton, MA, USA) and they were analyzed using the two-enzyme CK–LDH-assay (see Scheme 1A; the experimental details are given in the ESI†). After obtaining the optical responses (similar to those shown in Fig. 1 for the model solutions), we applied statistical analysis to evaluate the assay sensitivity and specificity for discriminating the samples of different ethnic origins. Receiver operating characteristic (ROC) analysis[52] was used to evaluate the performance of the assay and the possibility to distinguish between the CA (14 samples) and AA (14 samples) groups of human serum. Using ROC analysis, the best threshold (above which the absorbance changes correspond to the AA group) that yielded the maximum accuracy was determined. The area under the ROC curve (AUC) is a single measure summarizing the overall accuracy of the test. It represents the probability that the diagnostic test will correctly distinguish between the CA and AA samples. The AUCs from empirical and smooth ROC curves,[53] which expectedly give consistent results in this case (and the corresponding 95% confidence intervals; CI), were estimated for the data obtained with the CK–LDH-assay. The AUC of the empirical ROC curve was estimated by the trapezoidal method of integration, and the corresponding 95% CI was estimated with the method described by DeLong *et al.*[54] Smoothed ROC curves were additionally estimated using a non-parametric method. The kernel density function[55] was used to fit a smooth ROC curve to the data points because this method is free of parametric assumptions.[56] This smoothed-curve method outperforms the competing methods when the assumption of a normal distribution is violated. The bandwidth of the kernel function was fixed using the robust method developed by Sheather and Jones.[57] The AUCs of smooth ROC curves were obtained with corresponding 95% CIs computed with 2000 stratified bootstrap replicates as described elsewhere.[58] The AUC was 0.82 (95% CI: 0.64–1.00) from the empirical ROC curve and 0.80 (95% CI: 0.63–0.96) from the smooth ROC curve, Fig. 6, which means that the diagnostic test has an 82% chance to differentiate between CA and AA human serum. An absorbance change of 0.103 had the best corresponding sensitivity–specificity pair (the best tradeoff between them), *i.e.* the most accurate cut-off point for discrimination between CA and AA serum samples (the red point in Fig. 6).

Conclusions

The study performed on the model solutions mimicking CK and LDH biomarker composition in CA and AA groups demonstrated statistically proven recognition of samples with different ethnic

origins. It should be noted that only the biocatalytic cascade utilizing both biomarkers CK–LDH allowed CA and AA differentiation, while the analysis of CK alone was not able to show a statistically meaningful difference between the CA and AA groups. The analytical results obtained for the model solutions were confirmed when real serum samples from donors with known ethnic origins were applied. The developed method was tested on dried and aged serum samples allowing recognition of their ethnic origin. This method represents the first attempt to develop an on-field rapid analysis of biological fluids for forensic applications based on multi-enzyme biocatalytic cascades. The approach borrowed from unconventional enzyme-based computing[20,21] and originally applied to biomedical analysis[25,26] demonstrates promising perspectives for novel forensic serology applications. Further analytical applications for rapid identification of gender, age and other personal characteristics are feasible and they are under investigation in our lab. New tools (*e.g.* analytical kits or paper strips) for the rapid identification of biofluid origins are expected based on the present research.

Acknowledgements

This work was supported by the NSF award # CBET-1066397. The authors are grateful for help and advice on mathematical calculations provided by J. Sedláček.

References

1 *Forensic Analysis on the Cutting Edge: New Methods for Trace Evidence Analysis*, ed. R. D. Blackledge, Wiley-Interscience, Hoboken NJ, 2007.
2 R. Li, *Forensic Biology*, CRC Press, Boca Raton, FL, 2008.
3 R. C. Shaler, Modern forensic biology, in *Forensic Science Handbook*, ed. R. Saferstein, Prentice Hall, Upper Saddle River, NJ, 2002, pp. 529–546.
4 *Forensic Chemistry Handbook*, ed. L. Kobilinsky, Wiley, Hoboken, NJ, 2011.
5 V. Sikirzhytski, A. Sikirzhytskaya and I. K. Lednev, *Appl. Spectrosc.*, 2011, **65**, 1223–1232.
6 K. Virkler and I. K. Lednev, *Forensic Sci. Int.*, 2009, **193**, 56–62.
7 K. Virkler and I. K. Lednev, *Forensic Sci. Int.*, 2009, **188**, 1–17.
8 J. H. An, K.-J. Shin, W. I. Yang and H. Y. Lee, *BMB Rep.*, 2012, **45**, 545–553.
9 R. E. Gaensslen, *Sourcebook in Forensic Serology, Immunology, and Biochemistry*, U.S. Department of Justice, Washington, DC, 1983.
10 K. M. Elkins, *Forensic DNA Biology: A Laboratory Manual*, Academic Press, Oxford, UK, 2012.
11 M. Bauer, *Forensic Sci. Int.: Genet.*, 2007, **1**, 69–74.
12 J. S. Bonifacino, E. C. Dell'Angelica and T. A. Springer, *Immunoprecipitation. Curr. Protoc. Mol. Biol.*, 2001, ch. 10, unit 10.16, pp. 1–29.
13 J. M. Butler, *Fundamentals of Forensic DNA Typing*, Academic Press, Burlington, MA, 2009.
14 *DNA Testing and Data Banking; Law, Crime and Law Enforcement*, ed. E. J. Donovan and R. G. Mccullen, Nova Science Pub Inc, 2012.
15 S. H. James, P. E. Kish and T. P. Sutton, *Principles of Bloodstain Pattern Analysis: Theory and Practice*, CRC Press, Boca Raton, FL, 2005.
16 A. Eckenrode, E. G. Bartick, S. Harvey, M. E. Vucelick, B. W. Wright and R. A. Huff, *Forensic Sci. Commun.*, 2001, vol. 3(4), published by FBI on-line: http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/oct2001/eknrode.htm.
17 K. M. Horsman, J. M. Bienvenue, K. R. Blasier and J. P. Landers, *J. Forensic Sci.*, 2007, **52**, 784–799.
18 P. Liu, S. H. I. Yeung, K. A. Crenshaw, C. A. Crouse, J. R. Scherer and R. A. Mathies, *Forensic Sci. Int.: Genet.*, 2008, **2**, 301–309.
19 *Molecular and Supramolecular Information Processing: From Molecular Switches to Logic Systems*, ed. E. Katz, Wiley-VCH, Weinheim, 2012.
20 *Biomolecular Information Processing – From Logic Systems to Smart Sensors and Actuators*, ed. E. Katz, Wiley-VCH, Weinheim, 2012.
21 E. Katz and V. Privman, *Chem. Soc. Rev.*, 2010, **39**, 1835–1857.
22 J. Zhou, J. Halámek, V. Bocharova, J. Wang and E. Katz, *Talanta*, 2011, **83**, 955–959.
23 J. Halámek, V. Bocharova, S. Chinnapareddy, J. R. Windmiller, G. Strack, M.-C. Chuang, J. Zhou, P. Santhosh, G. V. Ramirez, M. A. Arugula, J. Wang and E. Katz, *Mol. BioSyst.*, 2010, **6**, 2554–2560.
24 J. Halámek, J. R. Windmiller, J. Zhou, M.-C. Chuang, P. Santhosh, G. Strack, M. A. Arugula, S. Chinnapareddy, V. Bocharova, J. Wang and E. Katz, *Analyst*, 2010, **135**, 2249–2259.
25 J. Wang and E. Katz, *Anal. Bioanal. Chem.*, 2010, **398**, 1591–1603.
26 E. Katz, J. Wang, M. Privman and J. Halámek, *Anal. Chem.*, 2012, **84**, 5463–5469.
27 L. Halámková, J. Halámek, V. Bocharova, S. Wolf, K. E. Mulier, G. Beilman, J. Wang and E. Katz, *Analyst*, 2012, **137**, 1768–1770.
28 H.-G. Xie, R. B. Kim, A. J. J. Wood and C. M. Stein, *Annu. Rev. Pharmacol. Toxicol.*, 2001, **41**, 815–850.
29 W. Kalow, *Clin. Pharmacokinet.*, 1982, **7**, 373–400.
30 J. Lovric, *Introducing Proteomics: From Concepts to Sample Separation, Mass Spectrometry and Data Analysis*, Wiley, Hoboken, NJ, 2011.
31 P. A. Deuster, F. G. O'Connor, K. Kenney, Y. Heled, S. Muldoon, C. Contreras-Sesvold and W. W. Campbell, *Creatine Kinase Clinical Considerations: Ethnicity, Gender and Genetics, Abstract at the conference: RTO-MP-HFM-181-Human Performance Enhancement for NATO Military Operations (Science, Technology and Ethics)*, http://www.cso.nato.int/pubs/rdp.asp?RDP=RTO-MP-HFM-181.
32 N. Samadi, G. S. Cembrowski and J. Chan, *Clin. Biochem.*, 2007, **40**, 206–212.
33 S. Okinaka, H. Sugita, H. Momoi, Y. Toyokura, H. Kumagai, S. Ebashi and Y. Fujie, *Trans. Am. Neurol. Assoc.*, 1959, **84**, 62–64.
34 J. B. Henson and R. R. Rao, *Can. J. Comp. Med. Vet. Sci.*, 1966, **30**, 157–159.

35 D. Kumbhare, W. Parkinson and B. Dunlop, *J. Spinal Disord. Tech.*, 2008, **21**, 49–54.
36 J. E. Olerud, L. D. Homer and H. W. Carroll, *Arch. Intern. Med.*, 1976, **136**, 692–697.
37 A. Kratz, M. Ferraro, P. M. Sluss and K. B. Lewandrowski, *N. Engl. J. Med.*, 2004, **351**, 1548–1563.
38 S. B. Rosalki and J. H. Wilkinson, *Nature*, 1960, **188**, 1110–1111.
39 L. Noda, T. Nihet and M. F. Morale, *J. Biol. Chem.*, 1960, **235**, 2830–2834.
40 G. Forster, E. Bernt and H. U. Bergmeyer, in *Methods of Enzymatic Analysis*, ed. H. U. Bergmeyer, Academic Press, Inc., NY, 2nd edn, 1974, vol. II, pp. 789–793.
41 http://www.sigmaaldrich.com/etc/medialib/docs/Sigma/General_Information/creatine_phosphokinase.Par.0001.File.tmp/creatine_phosphokinase.pdf.
42 A. H. B. Wu, *Tietz Clinical Guide to Laboratory Tests*, Saunders-Elsevier, St. Louis, MO, 4th edn, 2006.
43 R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012, ISBN 3-900051-07-0.
44 The R Project for Statistical Computing, http://www.R-project.org.
45 R. D. C. Hirata, M. H. Hirata, B. Strufaldl, R. A. Possik and M. Asai, *Clin. Chem.*, 1989, **35**, 1385–1389.
46 A. G. Rumley, A. R. Pettigrew, M. E. Colgan, R. Taylor, S. Grant, A. Manzie, I. Findlay, H. Dargie and A. Elliott, *Br. J. Sports Med.*, 1985, **19**, 152–155.
47 C. Walck, *Handbook on Statistical Distributions for Experimentalists*, University of Stockholm, Internal Report SUF-PFY/96-01, http://www.stat.rice.edu/~dobelman/textfiles/DistributionsHandbook.pdf.
48 M. Rosenblatt, *Ann. Math. Stat.*, 1956, **27**, 832–837.
49 E. Parzen, *Ann. Math. Stat.*, 1962, **33**, 1065–1076.
50 A. Basu, H. Shioya and C. Park, *Statistical Inference: The Minimum Distance Approach*, CRC Press, Boca Raton, 2011, ISBN 978-1-4200-9965-2.
51 M. M. Deza and E. Deza, *Dictionary of Distances*, Springer, Berlin/Heidelberg, 2009, ISBN 978-3-642-00233-5.
52 M. H. Zweig and G. Campbell, *Clin. Chem.*, 1993, **39**, 561–577.
53 J. E. Hanley, *Med. Decis. Making*, 1988, **8**, 197–203.
54 E. R. DeLong, D. M. DeLong and D. L. Clarke-Pearson, *Biometrics*, 1988, **44**, 837–845.
55 K. H. Zou, W. J. Hall and D. E. Shapiro, *Stat. Med.*, 1997, **16**, 2143–2156.
56 D. Faraggi and B. Raiser, *Stat. Med.*, 2002, **21**, 3093–3106.
57 S. J. Sheather and M. C. Jones, *J. Roy. Stat. Soc. B Stat. Meth.*, 1991, **53**, 683–690.
58 J. Carpenter and J. Bithell, *Stat. Med.*, 2000, **19**, 1141–1164.