Dengue is a very common viral disease in tropical countries. From time to time it has become epidemic, hindering economics and affecting the social system. Early starting treatment reduces recovery time and suffering of the patients. But, the usual diagnostics is done by detection of the antibodies, which can only be done after five days. Other parts of the metabolomics might change much faster and therefore the small molecule content in the blood might change much faster than that. By using NMR spectroscopy we attempted to differentiate blood plasma from infected patients from healthy subjects and afterwards identify the consistent changes. Several of these were found and are discussed.

Introduction

According to WHO records, 50-100 million people are infected with Dengue every year. Within 100 countries, about 550 thousand patients are hospitalized and about 20 thousand die every year. The early diagnosis of Dengue is done by eliminating other possible causes of the observed symptoms, a direct antibody diagnosis becomes possible only after several days into the disease. Hence, for the first days, a patient is either not treated (not strong enough symptoms) or treated possible only after several days into the disease. Hence, for the first days, a patient is either not treated (not strong enough symptoms) or treated “on suspicion”. A method of early identification for Dengue is need [1].

The blood is the main means of transport inside the human body, not only for proteins (like antibodies), but also for small molecules. Due to changes in their metabolism, infected cells could be producing different secondary metabolites, which will show up in the blood plasma. This changes should appear quickly, much faster than antibodies. Blood plasma has been extensively studied by NMR [2-4] and many small molecules can be associated to certain NMR signals. A lot of work has been done in order to characterize the lipids in plasma samples [5-8], but hardly have other molecules been looked at [9-11]. The quantitative nature of the proton NMR experiment allows the observation of relative changes in the concentration of the different metabolites, as changes in the NMR signals. With the help of Principal Component Analysis (PCA) the characterization of consistent changes becomes possible [4,9,11-17].

A series of consistent changes were assigned and the metabolic function of the identified molecules were investigated. Of course many changes are in general characteristic for infections, but there were a few exceptions, specially when looking the the combination of different factors. Hence, not only might we be able to differentiate the plasma samples of healthy and infected subjects, but also subjects with different infections.

Materials and Methods

In our study we have analyzed blood plasma from healthy and infected subjects using proton NMR. The samples originate from the collection from the “Instituto Oswaldo Cruz”, were collected under non-controlled conditions over several years and kept deep frozen. A volume of 0.5 ml of the samples was filtered with 3kDa filters (Millipore Amicon Ultra-4 Centrifugal Filter Units 3kDa, centrifuged at 4000g), freeze-dried and re-suspended in 0.5 ml D2O. 1-D NOE spectra [9,16] were acquired on a Bruker Avance 500 MHz with BBO probe, using a...
D1 of 2s, 64 scans and pre-saturation on the water signal. The spectra were calibrated on the glucose methyl at 1.33 ppm and then subjected to an extensive Principal Components Analysis (PCA) [11,16-18] using the software Amix version 3.8.6, in order to obtain a differentiation. Signal ranges that showed significant changes within the same group were excluded for the PCA, as well as the glucose signals.

All signals identified by the PCA were assigned to a metabolite via [3] and the observed change was noted. After a literature survey regarding the concentration change in the metabolite, a further analysis of the combination of different factors was attempted.

Results

In total 63 plasma samples were prepared as described and the obtained NMR spectra were evaluated. As it turns out, the width of the glucose peaks (3.5 ppm - 4.3 ppm) varies dramatically, and decided whether a spectrum could be used in the PCA or not, as shown in figure 1. After this evaluation only 51 plasma sample spectra could be used for further analysis. A total of 19 spectra, 5 being of healthy subjects, were randomly chosen and used to create the PCA model. The remaining 32 spectra, including 4 spectra of healthy subjects, were used as test-set.
The analysis of the plots reveals that the result is seemingly independent of the bucket width, the separation is always maintained. A series of different PCA models were created, varying several of the available parameters. The differentiation is very consistent and survives even dramatic parameter changes. For three different bucket widths (0.1 ppm, 0.05 ppm and 0.01 ppm) we obtained similar results, with PC2 vs PC1, PC2 vs PC3 and PC2 vs PC4 showing nicely separated data points (see figure 2). The plots for the different bucket sizes are composed of several pieces: A) The confidence plot, showing if all samples are within the model. The further the points are in the lower left quadrant, the better the model describes them. B) The deviation of each sample from the model. It shows how well the model really represents an average over all samples. A sample with distance zero from the model would be located in the middle of the plot, points above the middle have a positive deviation, points below negative. C) Number of PCs for the explanation of >95% of the data’s variation. D) Plot of different PCs against each other, showing the separation of the data. E) Plot of the buckets that are responsible for the positioning of the samples in the previous plots. The most important buckets can be identified by comparing the plots D and E for the same PCs.

Figure 4: Superposition of randomly chosen spectra (healthy subjects in red and blue, infected subjects in green and violet) showing the difference as identified by the PCA at 1.86 ppm (citrulline), 1.92 ppm (lisine, arginine and acetate), 2.04 ppm (glycoproteins), 2.05 ppm (proline), 2.08/2.08 ppm (glutamine), 2.13 ppm (metionine), 2.15 ppm (glutamate) and 2.24 ppm (valine).

Figure 5: Superposition of randomly chosen spectra (healthy subjects in red and blue, infected subjects in green and violet) showing the difference as identified by the PCA at 3.16 ppm (phenylalanine), 3.24 ppm (arginine), 3.35 ppm (proline), 3.95 ppm (glycerol), 3.97 ppm (phenylalanine) and 3.98 ppm (histidine).

A series of different PCA models were created, varying several of the available parameters. The differentiation is very consistent and survives even dramatic parameter changes. For three different bucket widths (0.1 ppm, 0.05 ppm and 0.01 ppm) we obtained similar results, with PC2 vs PC1, PC2 vs PC3 and PC2 vs PC4 showing nicely separated data points (see figure 2). The plots for the different bucket sizes are composed of several pieces: A) The confidence plot, showing if all samples are within the model. The further the points are in the lower left quadrant, the better the model describes them. B) The deviation of each sample from the model. It shows how well the model really represents an average over all samples. A sample with distance zero from the model would be located in the middle of the plot, points above the middle have a positive deviation, points below negative. C) Number of PCs for the explanation of >95% of the data’s variation. D) Plot of different PCs against each other, showing the separation of the data. E) Plot of the buckets that are responsible for the positioning of the samples in the previous plots. The most important buckets can be identified by comparing the plots D and E for the same PCs.

| δ [ppm] | Metabolite Assignment | Change |
|---------|-----------------------|--------|
| 1.86    | citrulline            | ↓      |
| 1.91    | lisine, arginine, acetate | ↑  |
| 2.04    | glycoproteins         | ↓      |
| 2.05    | proline               | ↓      |
| 2.08/2.09 | glutamine            | ↑      |
| 2.13    | metionine             | ↑      |
| 2.15    | glutamate             | ↓      |
| 2.24    | valine                | ↓      |
| 3.24    | arginine              | ↑      |
| 3.16    | phenylalanine         | ↓      |
| 3.35    | proline               | ↓      |
| 3.95    | glycerol              | ↓      |
| 3.97    | phenylalanine         | ↓      |
| 3.98    | histidine             | ↑      |

Table 1: Changes in signal intensities observed between plasma samples of healthy and Dengue infected subjects and corresponding assignment to metabolites.

The analysis of the plots reveals that the result is seemingly independent of the bucket width, the separation is always maintained.
But the narrowest bucket width shows the best model, which can be seen in the distance plot B. It also becomes very clear that only 8 buckets are relevant for the separation of the data, consistent for all models. Therefore all models were used for the test-set analysis. Figure 3 shows the result of our PCA model applied to the classification of the test-set samples. As before, the classification works better when the buckets become narrower. The model with the bucket width 0.01 ppm only has one sample classified erroneously, but it also was classified as being out of the model's confidence interval.

Overall PC2 is the most important component for the differentiation between healthy and infected subjects, whereas PC1, PC3 and PC4 contribute much less to the differentiation. From the PCA model we can identify the most important regions responsible for that result, mainly around 1.92 ppm (Figure 4) and 3.38 ppm (Figure 5). According to assignments published in the literature for blood plasma [3] we can associate the varying signals to a series of compounds: proline, lysine, arginine, acetate, treonine, β-glucose, pyruvate, citroline, glutamine, isoleucine and others. Table 1 shows a summary of all changes that could be identified in this ranges.

Discussion and Conclusion

The increased level of phenylalanine can be correlated to the the immune activation marker neopterin. The oxidative stress due to immune activation destroys the cofactor needed for phenylalanine (4-)–hydroxylase, hence the serum concentrations of phenylalanine increases [19]. The observed decrease of citrulline level and increase of proline level combined with unchanged arginine levels can be associated to certain production paths of nitric oxide, common response to viral infections. All three compounds are connected in the metabolism. Arginase hydrolyses L-arginine to urea and L-ornithine which, in turn, is a precursor for the synthesis L-proline by the enzyme ornithine aminotransferase. Alternatively arginine can be metabolized by nitric oxide synthase (NOS) to produce nitric oxide (NO) and L-citrulline. It has been shown that Dengue patients have increased NO levels [20]. Frequently Dengue is compared to Sepsis, as one can lead to the other [21]. It has been shown that Sepsis patients have higher protein breakdown levels [22,23], which are the major source of L-arginine in the metabolism, and therefore compensating the NOS arginine use, whilst maintaining a low citrulline level.

Overall, not all differences found in the NMR spectra could be associated to known metabolic effects. Further investigations will be carried out on them. But, for the moment, the observed level decreases of citrulline and proline combined with a constant arginine level seem to be a possible way to identify Dengue.

Acknowledgments

The authors wish to acknowledge funding for this project from FAPERJ - Edital 10/2008 and CAPES.

References

1. (1997) Dengue haemorrhagic fever: diagnosis, treatment, prevention and control. (2ndedn), Geneva: World Health Organization, USA.
2. Wewers RA, Engelke U, Heerschap A (1994) High-Resolution 1H-NMR Spectroscopy of Blood-Plasma for Metabolic Studies. Clin Chem 40: 1245–1250.
3. Ala-Korpela M (1995) 1H NMR spectroscopy of Human Blood Plasma. Prol Nucl Magn Reson Spectrosc 27: 475–554.