International Journal of Biochemistry Research & Review

18(3): 1-8, 2017; Article no.IJBCRR.35595
ISSN: 2231-086X, NLM ID: 101654445

Analysis of Isotopic Peaks in Hydrogen/Deuterium Exchange Mass Spectra for Precise Measurement of Isotope Average Centroid Mass of Peptides and Proteins Using R Programming Language

Arumugaperumal Arun[1], Sreekala Narayanan[1], Subhasis Sarangi[1] and Amit Kumar Mandal[1*]

[1]Clinical Proteomics Unit, Division of Molecular Medicine, St. John's Research Institute, St. John's National Academy of Health Sciences, 100 Feet Road, Koramangala, Bangalore 560034, India.

Authors' contributions

This work was carried out in collaboration between all authors. Author AA analyzed the data, designed the programming software and wrote the manuscript. Author SN performed the experiment. Author SS helped in designing the programming software and interpreting the data. Author AKM designed the research project, interpreted the data and wrote the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/IJBCRR/2017/35595
Editor(s):
(1) Hector M. Mora Montes, Department of Biology, University of Guanajuato, Mexico.
Reviewers:
(1) Roger Mafua, University Agostinho Neto, Angola.
(2) Ageng Trisna Surya Pradana Putra, State Islamic University of Sultan Maulana Hasanuddin Banten, Indonesia.
Complete Peer review History: http://www.sciencedomain.org/review-history/20432

Short Communication

Received 20th July 2017
Accepted 4th August 2017
Published 9th August 2017

ABSTRACT

In hydrogen/deuterium exchange mass spectra, the shift in isotopically distributed peaks with deuterium incorporation into peptides results in an increase in the isotope average centroid mass. Due to limitations in the resolution of instruments and data processing by softwares, the observed peaks might have obscurity in its maxima values that can affect the centroid mass calculation. In the present study, using Gaussian function, the exact maxima of individual isotopic peaks of a distribution was calculated. Using those maxima and our customized program, the centroid mass was calculated. We propose that our method of calculating average centroid mass is mathematically more appropriate.

**Corresponding author: E-mail: amit@sjri.res.in;*

Keywords: Hydrogen/deuterium exchange; isotope average centroid mass; Gaussian fitting; mass spectrometry; MALDI.

1. INTRODUCTION

Hydrogen/deuterium (H/D) exchange is a molecular phenomenon wherein polar hydrogen atoms of proteins get exchanged with deuterium from polar solvent D_2O. H/D exchange monitored through mass spectrometry has emerged as a promising tool to explore conformational dynamics of proteins both *in vitro* and *in vivo* [1,2]. H/D exchange based mass spectrometry is used to understand the mechanism of protein-ligand interaction, protein folding and to establish bio-similar property of recombinant proteins used as drug molecules [3,4,5]. In continuous exchange method, the isotope exchange kinetics of a protein is quenched across varying time points and the exchanged protein is subjected to proteolytic digestion to analyse incorporation of isotopes using mass spectrometry platform.

Generally, 1% natural abundance of ^{13}C isotope results in the appearance of isotopically distributed peaks of natural peptides differing in mass by 1 Da. In H/D exchange kinetics, the replacement of each hydrogen with deuterium also leads to an increase in mass of molecules by 1 Da. With increase in number of deuterium incorporated, the envelope of isotopic distribution of peaks of a peptide shifts towards higher mass till the saturation is achieved. In the analysis of H/D exchange mass spectra, the isotope average centroid mass of the peptide is calculated. The change in the centroid mass across varying time points is translated in terms of the conformational dynamics of proteins. Thus, it is very important to calculate the centroid mass in the H/D exchange mass spectra of peptides with high precision. There are many softwares available for the calculation of centroid mass, namely HX-Express [6], DynamX, HD-examiner, Auto-HD [7], DEX [8], Hydra [9], Hexicon2 [10], Ex-MS [11], HDX-Workbench [12] and Mass Analyzer [13].

However, due to limitations in the resolution of mass spectrometers, sometimes it becomes difficult to find out the exact maxima of each of the isotopic peaks and the obtained centroid mass of the isotopic distribution might appear to be incorrect in its decimal place-value. In the present study, using R programming language, we developed a method of fitting of individual isotopic peaks to a Gaussian function to calculate peak maxima of the curve. Subsequently the isotope average centroid mass of the peptide was calculated in a precise manner.

2. MATERIALS AND METHODS

The mass spectra used for the analysis were obtained by the proteolytic digestion of oxy and deoxy forms of normal human hemoglobin (HbA) using pepsin, as described in Narayanan et. al. [2]. H/D exchange data of 14 peptic peptides (9 from β globin chain and 5 from α globin chain) of deoxy and oxy HbA across varying time points were taken from our previously published results [2].

2.1 Analysis

MALDI mass spectra of isotope exchanged peptides were used for the analysis. The spectrum was smoothed using 5x5 point Savitsky-Golay filter (MassLynx software, Waters UK). Monoisotopic mass is the sum of elemental masses of a molecule wherein all constituent elements are present as most abundant natural isotopes. The spectrum of peptides was selected manually. To obtain a rough estimate of the width of isotopic distribution, the monoisotopic mass of the molecular ion was divided by 110 Da, the average mass of an amino acid. The approximate number of exchangeable peptide backbone amide hydrogens was calculated excluding the contribution of α-amino group of N-terminal amino acid and proline residue of the peptide. The above mass window in the smoothed mass spectra, starting from monoisotopic peak to the end of predicted isotopic distribution, was copied onto an MS-Excel spread sheet. Programming language, R ver 3.3.2 [14] was used for implementing multiple steps of calculation of centroid mass. The data points stored in the excel sheet was read into the R software using xlsx package [15]. The decimal place corrected monoisotopic mass was calculated from the input monoisotopic mass value by searching a window of ±0.5 Da about it. In the present study, we used MALDI mass spectra of peptides where unipositively charged ions survive. Thus all isotopic peak values were extracted by repeating the search for every 1 Da intervals. The most intense peak closest to the monoisotopic peak was considered as the base peak of a distribution. To calculate isotope average centroid mass, peaks that constitute the isotopic distribution were selected by employing

the following constraints: (i) intensity of a peak in the distribution must be greater than or equal to 20% of base peak intensity, (ii) any peak in the isotopic distribution in the higher m/z side of base peak is lower in its intensity compared to the neighbouring peak in low m/z side. On the right side of base peak, the distribution was truncated when the successive peak intensity becomes higher than the previous peak. To reduce noise in the spectrum, a cut-off value of 20% of base peak intensity was introduced for the entire distribution. In the next step, every individual peaks in the isotopic distribution was fitted into a Gaussian function f(x):

$$f(x) = a \exp(-(x-b)^2/2c^2) \quad (1)$$

The maximum intensity of a peak, 'a', and the corresponding m/z value, 'b', were extracted from fitted curve. Nonlinear least square fitting consisting of Gauss-Newton algorithm was used for the fitting process. Two virtual points, one on each side of the distribution, at the start and at the end of the envelope at 20% of the base peak were calculated using linear interpolation between the apex of extreme peaks and first data point below the 20% cut-off line of the same peak. Intensity envelope of the entire isotopic distribution was rebuilt using the points obtained through Gaussian fitting and the virtual points. The centroid of the distribution was calculated as follows:

$$\text{Centroid mass} = \sum_i (m/z)_i I_i / \sum_i I_i \quad (2)$$

Where I_i and $(m/z)_i$ are the intensity and the corresponding mass-to-charge ratio of i^{th} isotopic peak in a distribution [6].

Using this method, the centroid masses were calculated for 14 peptides of oxy form of HbA. In order to compare our method with HX-Express software, centroid values of those peptides were recalculated using HX-Express. The theoretical isotope average mass for those peptides were calculated using MS-Product tool of Protein Prospector (http://prospector.ucsf.edu/prospector/mshome.htm). Statistical software Origin8 was used to perform ANOVA on the calculated values.

We have provided the program code and instructions to run the program in supplement 1. A representative peptide mass spectrum with monoisotopic mass 931.6 Da has been provided in supplement 2.

3. RESULTS AND DISCUSSION

Natural abundance of different isotopes of the major constituent elements of proteins is as follows: Carbon (^{12}C- 98.93%, ^{13}C- 1.07%, ^{14}C- <0.01%); Oxygen (^{16}O- 99.757%, ^{17}O- 0.038%, ^{18}O- 0.205%); Nitrogen (^{14}N- 99.636%, ^{15}N- 0.364%); Hydrogen (^{1}H- 99.9885%, ^{2}H- 0.0115%, ^{3}H- <0.01%) [16]. Thus, depending on the elemental composition of a molecule, the highest peak in the isotopic distribution, *i.e.,* the base peak, need not always be the monoisotopic peak. In general for proteolytic fragments of a protein, which spans the molecular mass range of 1 to 4 kDa, three different types of isotopic distributions are observed. For peptides with number of carbon atoms <100, the monoisotopic peak is the base peak, peptides with carbon atoms ≥100 to <200, the second isotopic peak is the base peak and peptides with carbon atoms ≥200 to <300, the third isotopic peak is the base peak.

In isotope exchanged peptides, the nature of distribution of isotopic peaks is dominated by the kinetics of incorporation of deuterium in the molecule. In our H/D exchange experiment, three types of isotopic distributions were observed in the proteolytic peptides of oxy HbA. In the first type, monoisotopic peak and base peak were identical. A peptic peptide of α globin chain of unexchanged oxy HbA with mass 1585.9 Da showed that the monoisotopic peak was the base peak (Fig. 1a). In the mass spectra, the monoisotopic peak had the intensity 320.8. 20% cut-off of the base peak intensity was 64.16. Peaks that were below this cut-off value were eliminated in the subsequent calculation. The remaining peaks were fitted using equation (1) and the exact maxima and the corresponding mass were extracted. The intensity envelope for centroid mass calculation was rebuilt using the extracted peak maxima values. Two virtual points [(1585.665 Da, I = 64.16) and (1588.961 Da, I = 64.16)] were also added on both sides of the envelope. According to equation (2), the calculated centroid mass was found to be 1586.989. Importance of fitting individual isotopic peaks can be exemplified by a case wherein two consecutive points within the monoisotopic peak of a peptide with mass 1308.7 Da, a β globin fragment, were found to have same intensities [(1308.75 Da, I=10340) and (1308.76 Da, I=10340)]. The fitted curve of monoisotopic peak is shown in Fig. 1b. The Gaussian fitting in our method provided the point (1308.757 Da, I=10404) as the exact peak maxima.

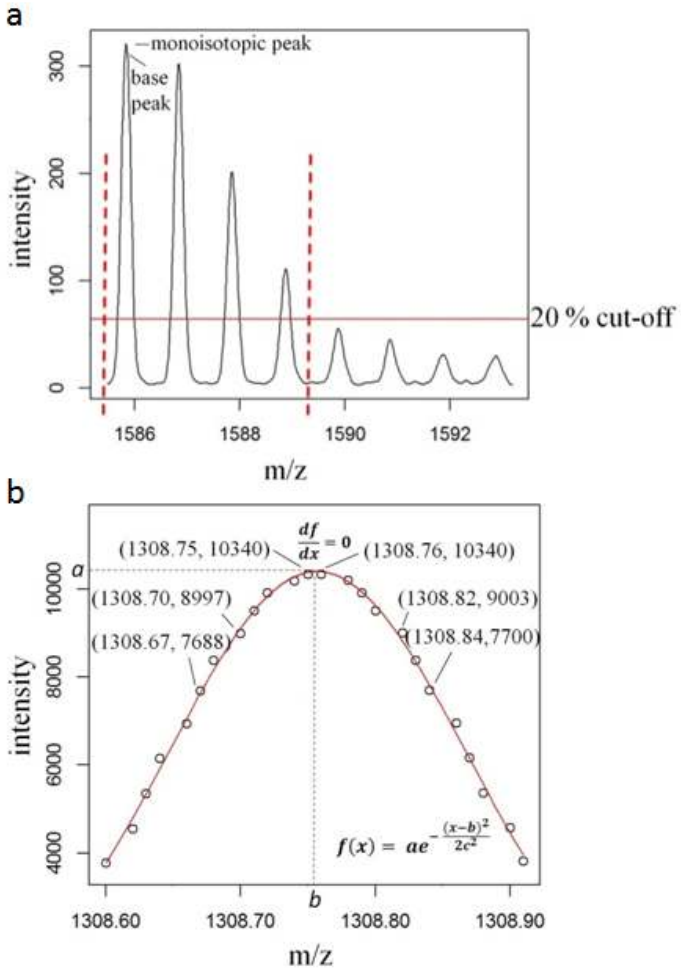

Fig. 1. (a) Representative mass spectra of isotopic peak distribution of a peptic peptide of hemoglobin. (b) Fitting of an individual peak with Gaussian function to calculate maxima of the peak

In the second type of isotopic peak distribution, the observed intensity of monoisotopic peak became less than 20% of the base peak intensity. 6 hours time point of a α globin fragment of oxy HbA with mass 2051.1 Da is shown in Fig. 2a. The intensity of the monoisotopic peak (2051.0991 Da) was 44.52, whereas the base peak (2054.0972 Da) was 267.70. Thus, the monoisotopic peak intensity was lower than 20% base peak intensity (I=53.54) and subsequently it was discarded in the centroid mass calculation. The individual isotopic peaks were fitted using equation (1) and the exact peak maxima and corresponding mass were extracted. The intensity envelope was rebuilt using exact maxima values of isotopically distributed peaks. Two virtual points, at the beginning (2051.919 Da) and at the end (2058.173 Da) at 20% cut-off intensity (53.54) were added. Using equation (2), the calculated centroid mass was found to be 2054.755 Da.

In the third type, two successive isotopic peaks were found to have very close intensities resulting in difficulty to identify the base peak. The solution was obtained by the fitting of individual isotopic peaks using Gaussian function and extracting the exact maxima values of both peaks. For example, 140 min time point of deuterium exchanged fragment of deoxy HbA with mass 1585.9 Da is shown in Fig. 2b. Two consecutive peaks had same intensities [(1586.8552 Da, I=131.10) and (1587.8514 Da, I=131.10)]. After fitting the isotopic peaks in Gaussian function, the obtained maxima were (1586.855 Da, I=129.81677) and (1587.861 Da,

I=130.93720). Thus, our method of fitting of individual isotopic peaks of the distribution provided the actual base peak. While rebuilding the isotopic distribution envelope with the maxima of individual isotopic peaks, 20% cut-off of base peak intensity was applied. This correct truncation of isotopic distribution envelope is shown in Fig. 2b. It was clear that the next isotopic distribution started to emerge but those were not included in the envelope of peptide of interest.

H/D exchange kinetics was monitored for both deoxy and oxy states of HbA across varying time points. We used our customized program to calculate the isotope average centroid mass of 14 peptic peptides, obtained from α and β globin chains of HbA. Isotope average centroid mass of aforementioned peptides for three different time points (5 min, 30 min and 105 min) of both the oxy and deoxy states are shown in Table 1. The centroid mass of those peptides at other time points (0, 10, 20, 40, 50, 60, 75, 90, 120, 140, 160, 180 and 360 min) of deoxy and oxy states of HbA have been provided in supplement 3 and supplement 4 respectively. Supplement 5 provides table with the centroid values calculated for 14 unexchanged peptides of the oxy HbA using our method, HX Express and Protein Prospector. The theoretical average mass values calculated by Protein Prospector were considered as standard values of centroid mass of the experimental peptides. ANOVA analysis showed that there was no statistical significant difference (P = 1) between the values calculated by these three methods.

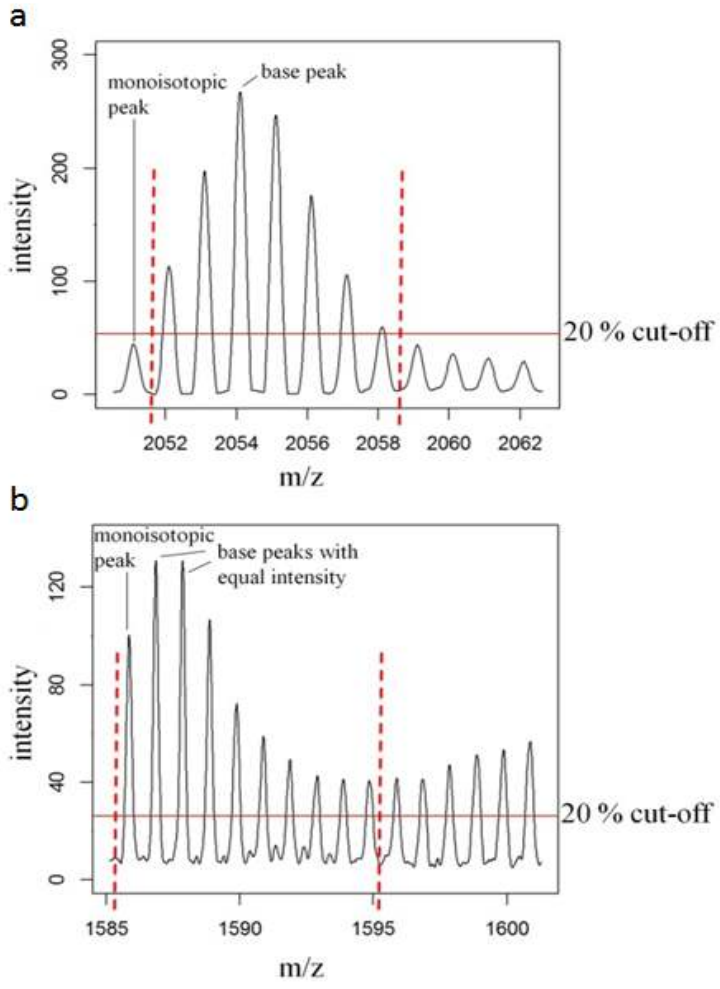

Fig. 2. (a) Representative mass spectra of isotopic distribution of peaks of peptic peptides of hemoglobin: (a) base peak is different from monoisotopic peak, (b) two consecutive peaks with equal intensities

Table 1. Centroid mass of H/D exchanged peptic peptides of HbA

Origin of peptide and amino acid sequence	deoxy HbA			oxy HbA		
	5 min	30 min	105 min	5 min	30 min	105 min
β130-146 YQKVVAGVANALAHKYH	1870.621	1871.033	1871.206	1870.628	1871.150	1871.229
β1-14 VHLTPEEKSAVTAL	1498.136	1499.242	1499.128	1497.154	1498.732	1499.589
α34-46 LSFPTTKTYFPHF	1587.656	1588.789	1589.512	1587.833	1589.570	1588.937
β86-102 ATLSELHCDKLHVDPEN	1923.137	1924.095	1925.354	1923.147	1926.075	1924.324
β32-41 LVVYPWTQRF	1309.741	1309.936	1310.654	1309.906	1310.437	1310.099
β32-40 LVVYPWTQR	1163.802	1164.159	1164.469	1163.143	1163.670	1165.060
α110-141 AAHLPAEFTPAVHASLDKFLASVSTVLTSKYR	3432.094	3433.24	3433.117	3431.430	3432.700	3433.843
β103-110 FRLLGNVL	932.212	932.253	932.437	931.962	932.304	932.326
β15-31 WGKVNVDEVGGEALGRL	1801.308	1802.297	1802.624	1800.732	1802.207	1802.779
α1-29 VLSPADKTNVKAAWGKVGAHAGEYGAEAL	2915.141	2916.386	2917.615	2914.328	2917.110	2917.019
β22-31 EVGGEALGRL	1002.697	1003.292	1003.327	1002.110	1003.197	1003.667
α1-32 VLSPADKTNVKAAWGKVGAHAGEYGAEALERM	3331.395	3332.624	3333.850	3330.785	3333.398	3333.151
α1-33 VLSPADKTNVKAAWGKVGAHAGEYGAEALERMF	3478.485	3479.749	3481.032	3477.920	3480.520	3480.301
β111-129 VCVLAHHFGKEFTPPVQAA	2052.318	2053.730	2053.934	2052.734	2053.679	2054.219

Adapted with permission from Analytical Chemistry 2015, 87 (23), 11812-11818. Copyright 2015 American Chemical Society

In the calculation of centroid mass of isotopic distribution of peaks in the mass spectra, proper selection of the distribution of peaks is crucial. Although the commonly used softwares such as HX Express assigns the isotopic distribution of peaks, the maxima value of each individual peak is taken from the maxima of peak list of recorded data by the mass spectrometer. In our method, fitting of individual isotopic peaks of the distribution using Gaussian function has been used to find out exact maxima of each peak. We propose that although there was no statistically significant difference between the centroid values calculated by the present method and by HX Express, our method of calculating maxima value of each peak is mathematically more appropriate and it avoids any errors incorporated due to instrument bias and/or post-recording data analysis.

4. CONCLUSIONS

Gaussian fitting to determine the exact peak maxima for individual peaks in an isotopic distribution and subsequent calculation of the average centroid mass is mathematically more appropriate approach for H/DX data analysis.

SUPPLEMENTARY DATA

Supplement 1 has the program code and step by step instructions to run the program. Supplement 2 contains a representative peptide mass spectrum with monoisotopic mass 931.6 Da. Supplement 3 contains a table with the centroid mass of H/D exchanged peptic peptides of deoxy HbA at various time points. Supplement 4 contains a table with the centroid mass of H/D exchanged peptic peptides of oxy HbA at various time points. Supplement 5 contains a table comparing centroid mass calculated by three softwares for an isotopic distribution.

ACKNOWLEDGEMENTS

Dr. Monita Muralidharan is acknowledged for providing mass spectrometric data and help in analysis. Department of Biotechnology, India is acknowledged for providing financial support to Arun A in the form of JRF (DBT/2015/SJRI/447).

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Harrison RA, Engen JR. Conformational insight into multi-protein signaling assemblies by hydrogen–deuterium exchange mass spectrometry. Curr. Opin. Struct. Biol. 2016;41:187-93.
2. Narayanan S, Mitra G, Muralidharan M, Mathew B, Mandal AK. Protein Structure–Function Correlation in Living Human Red Blood Cells Probed by Isotope Exchange-based Mass Spectrometry. Anal. Chem. 2015;87(23):11812-8.
3. Chalmers MJ, Busby SA, Pascal BD, He Y, Hendrickson CL, Marshall AG, et. al. Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. Anal. Chem. 2006; 78(4):1005-1014.
4. Hu W, Walters BT, Kan ZY, Mayne L, Rosen LE, Marqusee S, et. al. Stepwise protein folding at near amino acid resolution by hydrogen exchange and mass spectrometry. Proc. Natl. Acad. Sci. USA. 2013;110(19):7684-7689.
5. Pan J, Zhang S, Borchers CH. Comparative higher-order structure analysis of antibody biosimilars using combined bottom-up and top-down hydrogen-deuterium exchange mass spectrometry. Biochim. Biophys. Acta. 2016;1864(12):1801-1808.
6. Weis DD, Engen JR, Kass IJ. Semi-automated data processing of hydrogen exchange mass spectra using HX-Express. J. Am. Soc. Mass Spectrom. 2006;17(12): 1700-3.
7. Palmblad M, Buijs J, Håkansson P. Automatic analysis of hydrogen/deuterium exchange mass spectra of peptides and proteins using calculations of isotopic distributions. J. Am. Soc. Mass Spectrom. 2001;12(11):1153-62.
8. Hotchko M, Anand GS, Komives EA, Ten Eyck LF. Automated extraction of backbone deuteration levels from amide H/2H mass spectrometry experiments. Protein Sci. 2006;15(3):583-601.
9. Slysz GW, Baker CA, Bozsa BM, Dang A, Percy AJ, Bennett M, et. al. Hydra: software for tailored processing of H/D exchange data from MS or tandem MS analyses. BMC Bioinformatics. 2009;10: 162.
10. Lindner R, Lou X, Reinstein J, Shoeman RL, F. Hamprecht A, Winkler A. Hexicon 2: automated processing of hydrogen-

deuterium exchange mass spectrometry data with improved deuteration distribution estimation. J. Am. Soc. Mass Spectrom. 2014;25(6):1018-28.
11. Kan ZY, Mayne L, Chetty PS, Englander SW. ExMS: Data analysis for HX-MS experiments. J. Am. Soc. Mass Spectrom. 2011;22(11):1906-15.
12. Pascal BD, Willis S, Lauer JL, Landgraf RR, West GM, Marciano D, et al. HDX workbench: Software for the analysis of H/D exchange MS data. J. Am. Soc. Mass Spectrom. 2012;23(9):1512-21.
13. Zhang Z, Zhang A, Xiao G. Improved protein hydrogen/deuterium exchange mass spectrometry platform with fully automated data processing. Anal. Chem. 2012;84(11):4942-9.
14. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2015. Available:URL https://www.R-project.org/
15. Adrian A. Dragulescu. xlsx: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files. R package version 0.5.7; 2014. Available:https://CRAN.R-project.org/package=xlsx
16. Haynes WM, Editor. CRC handbook of chemistry and physics. CRC Press; 2014.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/20432