Asian Journal of Chemistry; Vol. 25, No. 3 (2013), 1461-1466

ASIAN JOURNAL OF CHEMISTRY

http://dx.doi.org/10.14233/ajchem.2013.13074

Multivariate Curve Resolution of *p*-Coumaric Acid *o*-Hydroxylation Reaction Mechanism Catalyzed by Tyrosinase

M. SIRATI SABET[1], N. GHEIBI[2,*] and A. AHMADI[3]

[1]Department of Biochemistry and Genetic, Qazvin University of Medical Sciences, Qazvin, Iran
[2]Cellular and Molecular Research Center, Qazvin University of Medical Sciences, Qazvin, Iran
[3]Department of Biology, Science Faculty, Science and Research Branch of the Islamic Azad University, Tehran, Iran

*Corresponding author: E-mail: gheibi_n@yahoo.com

(*Received*: 16 November 2011; *Accepted*: 12 September 2012) AJC-12138

Multivariate curve resolution is able to extract the number of components involved in a complex spectral feature and attribute the resulting spectra to chemical compounds. The quantification of the individual spectral contributions develop kinetic models for the process with or without a prior knowledge. Data elaboration by multivariate curve resolution was able to explain the kinetics of degradation, giving the pure spectra of the involved species and the respective concentration variation. In this study, *p*-coumaric acid *ortho* hydroxylation of mushroom tyrosinase investigated using UV-VIS absorption spectra and chemometric resolution techniques. From the spectra of the reaction between 200 and 600 nm, number of components and their mole fractions in the reaction solution were determined using chemometric resolution technique. After the determination of components, multivariate curve resolution-alternating least square by initial estimates of spectral profiles and proper constraints, was used to resolve the data matrix into pure concentration and spectral profiles. Using chemical rank determination and subspace comparison the chemical species or chemical rank of matrix obtained three. So, there are three species during monophenolase reaction of mushroom tyrosinase.

Key Words: Multivariate curve resolution-alternating least square, Mushroom tyrosinase, Monophenolase.

INTRODUCTION

Tyrosinase or polyphenol oxidase is a copper enzyme widely distributed throughout the phylogenetic scale. It catalyses the hydroxylation of monophenol to *o*-diphenols and their subsequent oxidation to *o*-quinones, in both cases by molecular oxygen[1,2]. The kinetic behaviour of tyrosinase is very complex due to the contemporaneous occurrence of the enzymatic oxidation of monophenol and *o*-diphenol to *o*-quinone, on the one hand and the coupled non-enzymatic reactions of *o*-quinone, on the other[3,4].

Chemometric methods like principal component analysis and partial least squares regression have been used successfully in many applications over the years. The technique is well established in many analytical fields, like UV/VIS, NIR, IR spectroscopy[5] gas and liquid chromatography[6] and manufacturing processes[7].

Much research has been performed on solving the mixture analysis problem and extracting real spectra and concentration profiles from overlapping spectral data without making any prior assumptions about the composition of the system. Several mixture analysis methods are known, like evolving factor analysis[8], fixed size moving window evolving factor analysis[9], target factor analysis[10], classical curve resolution[11], weighted curve resolution[12], multivariate curve resolution[13] and to a certain extent also techniques like parallel factor analysis[14]. Using these techniques, we can mathematically estimate the evolution of the chemical contributions over time for a specific experiment. The measured spectral information at different wavelengths reveals the specific absorption and also the morphologies of the compounds due to the scatter produced when particles or solid biomass is present. However, the resulting spectra are difficult to interpret and they lack specificity. This disadvantage can be solved using curve resolution methods using the following three objectives of this study. The first is resolve the number of chemical compounds and intermediates simultaneously present in the mixture from a complex spectral signature. The second identification of these species by transforming mathematical solutions to real spectra and increasing their specificity. The third to quantify each component and transfer this information to a kinetic model without any prior assumption or knowledge of the chemical model involved. Thus, we demonstrate the possibilities and power of this method to estimate and quantification of intermediates during the cresolase reaction that mushroom tyrosinase

is affected on *p*-coumaric acid at moderate temperatures at pH 6.8. Resolving the UV/VIS spectra by multivariate curve resolution yields concentration profiles for the kinetic process without the need for any prior knowledge about the chemical model. Alternating least square allows to extract the pure spectra and concentration of the components in a mixture from a set of spectra with different composition. Its application is particularly useful to evaluate the kinetics of a chemical process, allowing also to calculate the concentration profiles of all the involved species. This information can be used to study the mechanism of the reaction.

EXPERIMENTAL

Mushroom tyrosinase (MT: EC 1.14.18.1) were prepared as previously described[15]. *p*-Hydroxycinnamic acid (*p*-coumaric acid, λ_{max} = 288 nm, ε = 19400 M^{-1} cm^{-1}) was purchased from from Merck™. All other reagents were homemade analytical grade. The water used was re-distilled and ion-free. The buffer used throughout this research was 10 mM phosphate buffer solution, pH = 6.8 and the corresponding salts were obtained from Merck. All the experiments were carried out at 20 ºC and all solutions prepared in doubly distilled water. Freshly prepared enzyme and substrate solutions were used in this work.

UV-VIS spectrophotometry: *p*-Coumaric acid in variable concentrations is used in this study as a substrate model compound for mushroom tyrosinase monophenolase reaction. The kinetic assays of monophenolase activity were carried out using cary spectrophotometer, 100 bio model, with jacketed cell holders. As illustrated in Fig. 1 the UV-VIS spectra were recorded in the wavelength range between 200 and 600 nm, at the times 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 and 17 min of the reaction time.

Theoretical background

Subspace comparison: Subspace comparison compares two subspaces, each of which described by a set of orthonormal vectors selected by a method suitable for variable selection such as orthogonal projection approach[16], simple-to-use interactive self modeling mixture analysis[17,18] and principal component analysis[19]. Although the different methods select different key variables, if correct number of variables has been selected, the vectors selected by different methods span the same subspace of the full row or column space of the matrix. Suppose two subspaces are defined as A = {a1, a2,..., ak} and B = {b1, b2,..., bk}, where K is the number of key variables or factors selected by different variables selection methods and is the same for both matrices. The vectors in A and B are orthogonalized by the Schmidt procedure[20]. The next step is to calculate tr(K) as:

$$tr(K) = Trace(A^TBB^TA)$$

where, tr(K) varies between 0 and K. Subspace discrepancy function, D(K), is calculated as follows:

$$D(K) = K - tr(K)$$

D(K) is the measure of that part of the subspaces, which is in orthogonal complement of the other. This becomes zero when two subspaces are identical. Eigen values are calculated on (A^TB) matrix and are utilized as:

$$\sin^2(\nu_k) = 1 - g_k$$

where, $\sin^2(\nu_k)$ is the largest principal angle between subspaces that are selected by different variable selection methods and represent the degree of agreement between variables for determination of the number of chemical species and g_k is the eigen value. Shen *et al.*[21] compared D(K) and $\sin^2(\nu_k)$ as a measure of disagreement between the subspaces by plotting both values for K components. The number of factors becomes equal to the largest value of K with D(K), or $\sin^2(\nu_k)$ close to zero.

Multivariate curve resolution-alternating least square (MCR-ALS): Multivariate curve resolution-alternating least square[22] is a commonly used technique that can resolve multi-component mixtures into a simple model consisting of a composition-weighted sum of the signals of the pure compounds. A data matrix X of dimension M × N is considered, where M is the number of spectra and N is the number of variables (*e.g.*, wavelengths). The aim of any resolution method is the optimal decomposition of a data matrix D into the product of two small matrices C and S, which contain as much information as possible about the pure component spectra and their concentrations, respectively.

$$D = CS^T + E$$

where, D(r × c) is the original data matrix, C(r × n) and S^T(n×c) are the matrices containing the pure response profiles related to the data variation in the row direction and in the column direction, respectively and E(r × c) is the error matrix, *i.e.* the residual variation of the data set that is not related to any chemical contribution. Parameters r and c are the number of rows and the number of columns of the original data matrix, respectively and *n* is the number of chemical components in the mixture or process. C and S^T often refer to concentration profiles and spectra (hence their names), although resolution methods are proven to work in many other diverse problems[23,24]. From the early days in resolution research, the mathematical decomposition of a single data matrix, no matter the method used, is known to be subject to ambiguities[25]. This means that many sets of paired C and S^T-type matrices can reproduce the original data set with the same fit quality. In plain words, the correct reproduction of the original data matrix can be achieved by using response profiles differing in shape (rotational ambiguity) or in magnitude (intensity ambiguity) from the sought (true) ones[26].

Eqn. (1) assumes an additive linear model, which should be valid for spectroscopic absorbance measurements when Beer's law applies. D contains the original absorbance spectra, very often as a function of time for a reaction process. C contains columns with the concentrations of the compounds involved in the data set, which are the kinetic profiles if the spectra describe a kinetic reaction. The rows of S contain the spectra for the compounds. E is the matrix of spectral residuals. Resolution of the peaks by using multivariate curve resolution-alternating least square for data matrix X consisted of the following steps: (i) use of principal component analysis-based methods such as evolving factor analysis[27] or purity-based methods such as subspace comparison[28] to determine the number of components; (ii) obtaining initial estimates of concentration profiles of spectra by evolving factor analysis, simple-to-use interactive self modeling mixture analysis and

orthogonal projection approach; (iii) use of alternating least squares on initial estimates to resolve the data matrix X into the pure component spectra and their related individual concentration profiles. The procedure starts by calculating the spectra by using the least square:

$$S = (X^TC)(C^TC)^{-1}$$

After this, a new set of concentration profiles (matrix C) can be obtained by least squares and refined using constraints:

$$C = (X^TS)(S^TS)^{-1}$$

Possible constraints are unimodality of concentration profiles non-negativity of the spectra and fixing a spectrum that is known and available during the alternating least square optimization process. Then sum of squares of the residuals, SSR, can be calculated using the following equations:

$$R = X - CS^T$$

$$SSR = \sum_{i=1}^{m}\sum_{j=1}^{n} r_{ij}^2$$

where, r_{ij} is the element of residual matrix. This procedure is repeated until the relative differences in the SSR values of two consecutive iterations are lower than a pre-defined convergence limit. At the end of the algorithm concentration profiles (C) and pure spectra (S) are stable and the data matrix was resolved.

Data analysis: The UV-visible spectra data collected in data matrix, D. This data matrix subjected to the multivariate curve resolution analysis. Then resulted concentration and spectra profiles of the components. The multivariate curve resolution calculations were performed using the Matlab toolbox multivariate curve resolution developed by R. Tauler and A. de Juan from the University of Barcelona, Spain (see http://www.ub.es/gesq/mcr/mcr.htm).

RESULTS AND DISCUSSION

The spectral information must be extracted and transferred into meaningful chemical assignments. Multivariate curve resolution is ideally suited to this task, since it provides information on the chemical constituents and their quantitative kinetic behaviour during processing without any prior knowledge. Although UV-VIS spectra are less informative in terms of chemical identification, they are very sensitive to changes in concentration. To demonstrate the ability of multivariate curve resolution to calculate concentration profiles during reactions from the measured spectra, *p*-coumaric acid used in this study as a substrate model compound for monophenolase reaction of mushroom tyrosinase.

Figs. 1a-d show the results of the reaction in which *p*-coumaric acid is affected by tyrosinase at 25 ºC and pH 6.8 under different concentration of substrate on exposure time of 17 min. While the *p*-coumaric acid is degraded, an intermediate state is formed, which diminishes when the final product is achieved. This signature can be attributed to consecutive reactions.

Chemometric analysis: The UV-visible spectra were used as experimental data for the chemometric analysis to determine the number of total components and their mole fraction in the reaction mixture during the monophenolase reaction of mushroom tyrosinase and to resolve the concentration and spectral profiles of these components. A critical step of curve resolution is determination of the number of components for monophenolase reaction of mushroom tyrosinase.

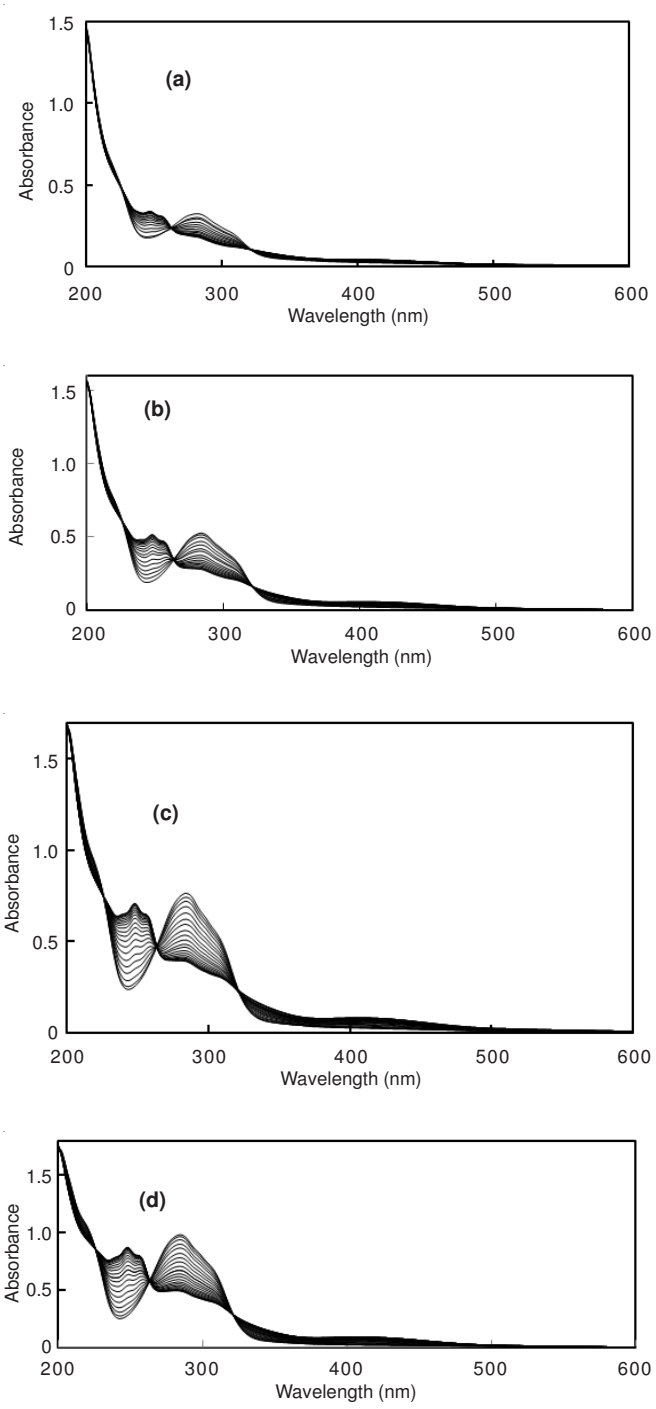


Fig.1. UV-visible spectra of cresolase reaction in variable concentrations of substrate (*p*-coumaric acid)

There are several methods for chemical rank determination in two-way data. In the present contribution we have applied subspace comparison[28] as a purity-based method for this purpose. In this method key spectra instead of full rank matrix are analyzed. The main idea behind subspace comparison is to extract information from the combinations of independent chemometric investigations. However, this method relies on other methods for variable selection, but it is simple to implement and fast to use. The results of this method for a

sample matrix (such as monophenolase reaction of mushroom tyrosinase) are shown in Fig. 2a and b.

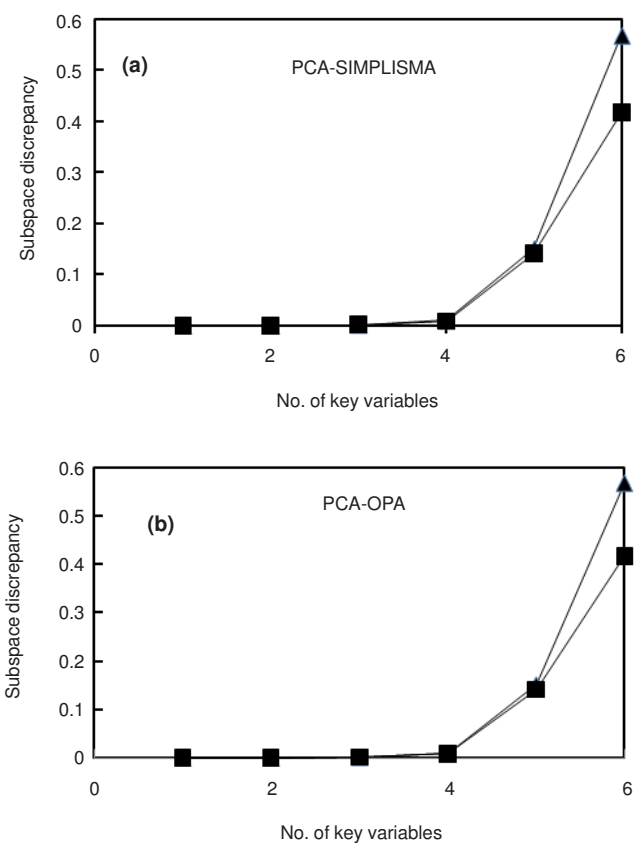


Fig. 2. Subspace plot for a sample matrix of monophenolase reaction of mushroom tyrosinase solution as a typical. (a) Comparison between simple-to-use interactive selfmodeling mixture analysis (SIMPLISMA) variables with principal component analysis (PCA) variables and (b) comparison between OPA variables and PCA scores. Subspace discrepancy functionD(K) (▲) and $\sin^2(n_k)$ (■)

In a subspace plot the number of chemical species is determined as the values for D(K) and $\sin^2(\nu_k)$ were equal to each other and close to zero. Fig. 2(a) compares the simple-to-use interactive self modeling mixture analysis variables with the orthogonal projection approach variables. Comparison between orthogonal projection approach variables and principal component analysis scores are shown in Fig. 2(b). These plots indicate that there are three variables by D(K) and $\sin^2(\nu_k)$ equal to each other and close to zero. Hence, the chemical rank of this matrix is three. In other word, there are three species in during the monophenolase reaction of mushroom tyrosinase. After the determination of components in each concentration of substrate, the two-dimensional data matrix has been uniquely resolved into concentration and spectral profiles of related components. For resolving the data, multivariate curve resolution-alternating least square method is applied. The initial estimates of spectral profiles are determined using evolving factor analysis, simple-to-use interactive self modeling mixture analysis and orthogonal projection approach methods. Fig. 3 (a-d) gives a resolved concentrations of components observed during the monophenolase reaction of mushroom tyrosinase at pH 6.8 and *p*-coumaric acid in variable concentrations. Fig. 4 (a-d) shows the corresponding spectra resolved by multivariate curve resolution-alternating least square. It should be pointed out that the initial estimate of spectral profile in this case was determined using orthogonal projection approach method. Other methods such as evolving factor analysis and simple-to-use interactive self modeling mixture analysis were also checked but their results were not appropriate.

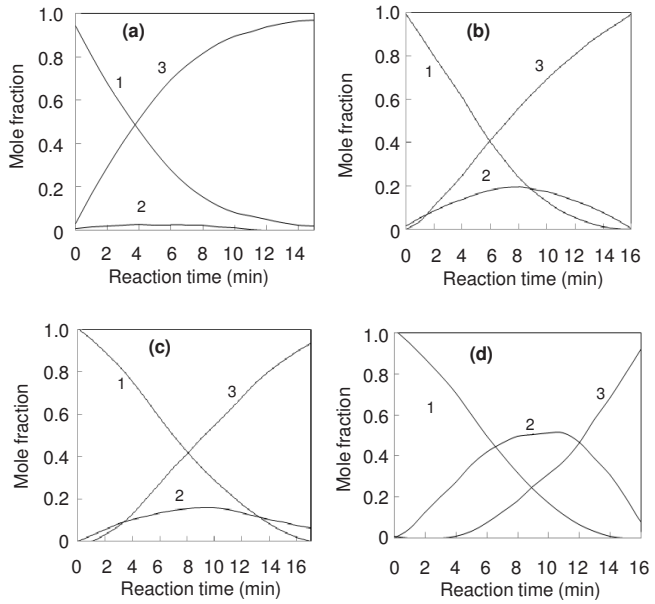
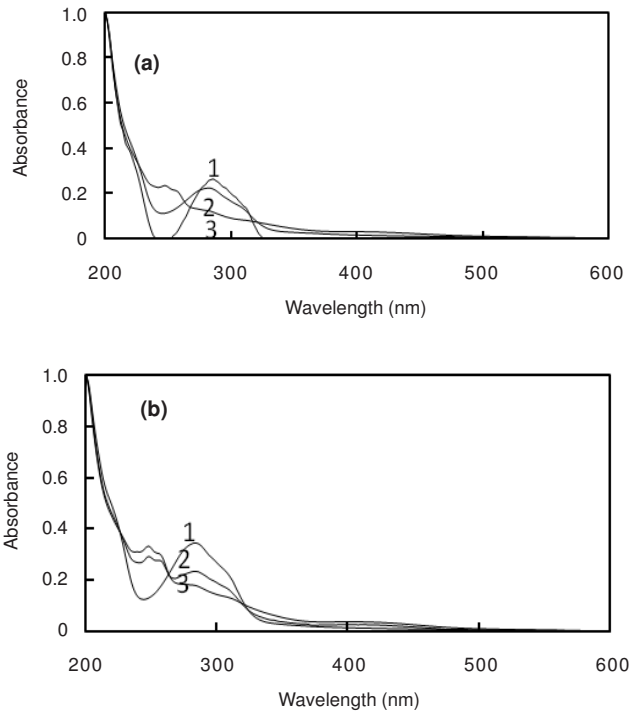


Fig. 3. Concentration profiles resolved by MCR-ALS that mole fraction of components (MFC) present in the monophenolase reaction of mushroom tyrosinase solution. The reaction was done in 10 mM PBS, pH 6.8, at 20 ºC and 70.8 µM enzyme. concentrations of *p*-coumaric acid (substrate): 10 µM (a), 20 µM (b), 30 µM (c), 40 µM (d), curve 1: *p*-coumaric acid (substrate), curve 2: intermediate and curve 3: end product



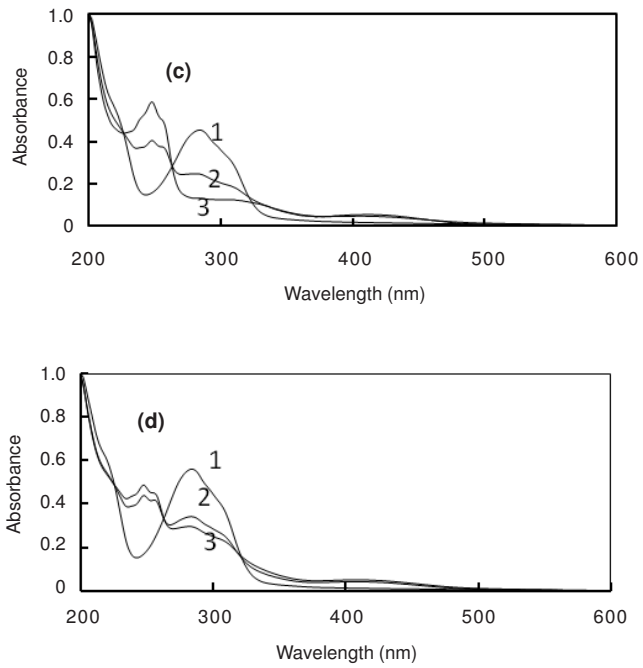


Fig. 4. Spectral profile of components resolved by MCR-ALS, in the monophenolase reaction of mushroom tyrosinase. The reaction was done in 10 mM PBS, pH 6.8, at 20 ºC and 70.8 µM enzyme. concentrations of *p*-coumaric acid (substrate) : 10 µM (a), 20 µM (b), 30 µM (c), 40 µM (d), curve 1: *p*-coumaric acid (substrate), curve 2: intermediate and curve 3: end product

The concentration profiles suggest that immediately after starting the reaction, the substrate curve goes to near zero, while a charge-transfer complex (electron-donor-acceptor complex) is formed and reaches its maximum value. Shortly after being at its peak, the charge-transfer complex curve goes down and converts into product (P). Moreover, concentration profiles obtained by soft modeling can be useful for formulating a suitable reaction model and supporting the model chosen[29]. The intermediate product starts converts into the final product. The proposed mechanism might be:

$$S + E \xrightarrow{K_1} [CTC] \xrightarrow{K_2} E + P$$

where, S is the substrate, here it denotes the *p*-coumaric acid, E the enzyme tyrosinase, [CTC] a charge-transfer complex that is an intermediate product and P the final product of the reaction.

Based on the crystal structure of tyrosinase obtained[4], we can describe the tyrosinase-specific catalytic mechanism (Fig. 5). At first, a peroxide ion, which forms a bridge with two Cu(II) ions in the oxy form of tyrosinase, acts as a catalytic base. As a result, a proton is abstracted from the phenolic hydroxyl. Subsequently, the deprotonated oxygen atom of monophenol binds to CuB at the sixth coordination site. At this time, CuB is hexa-coordinated by a tetragonal bipyramidal cage and an *ortho*-carbon of the substrate approaches the peroxide ion. One of two peroxide oxygens is then added to the *ortho*-carbon of monophenol. This monooxygenase reaction would be accelerated by the formation of a stable intermediate, in which newly generated oxygen atoms of diphenol bind to CuA. To form this state, His54, which is an axial ligand to CuA, must be released from the current position. This assumption is derived from the flexible feature of the residue His54 in the copper-free and Cu(II)-bound oxy forms. Simultaneously, His54 can act as a catalytic base for the deprotonation from the substrate. The resulting intermediate has the advantage of easy translation of electrons, resulting in the formation of the deoxy form of tyrosinase and quinone. Our proposed scheme does not fit the case of catechol oxidase because the bidentate intermediate cannot be formed because of the fixed conformation of His109, which corresponds to His54 of tyrosinase and the presence of the Phe261 lying just above the CuA site, which is vacant in the tyrosinase structure (Fig. 6)[30]. There is a consensus that the oxy form of tyrosinase can catalyze both the monooxygenase and oxidase reactions, whereas the met form lacks the monooxygenase activity. This can be explained as follows: the bidentate intermediate formation, which is permitted only by the oxy form, is essential for the monooxygenase reaction, but the mono-dentate intermediate, which is formed by the met form, is sufficient for an oxidase reaction, as proposed for catechol oxidase. Therefore, some compounds that bind to two Cu(II) ions in the bidentate form might be potent inhibitors of tyrosinase[4,30].

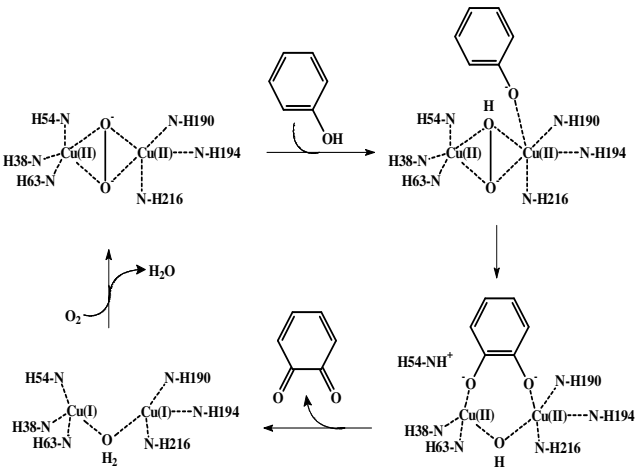


Fig. 5. Structure-based catalytic mechanism of tyrosinase. The oxy form of tyrosinase catalyzes the conversion of paracomaric acid (monophenol) to the corresponding quinone through the *ortho*-diphenol formation. In this scheme, His54 is released from the CuA site, resulting in the formation of the bidentate intermediate. The met and oxy forms of tyrosinase can catalyze the conversion of orthodiphenol to the corresponding quinone. This reaction should progress similarly to that of catecholoxidase

The use of multivariate resolution methods is an appropriate combination to study and interpret UV-visible spectral data chemical and biochemical reactions. The mechanism of the process and the nature of all the species involved can be known. The advantage of the above method compared to the classical methods is that it does not rely on the initial proposal of a specific kinetic model but estimates directly the changes in concentration, extracts the number of an analyte present and calculates pure component spectra, which are chemically meaningful. This method allows the deduction of reasonable mechanism of a reaction. In the *p*-coumaric acid *ortho* hydroxylation reaction by mushroom tyrosinase the chemical rank determinations and subspace comparisons showed three species during monophenolase reation of mushroom tyrosinase that correspond with crystallographic data.

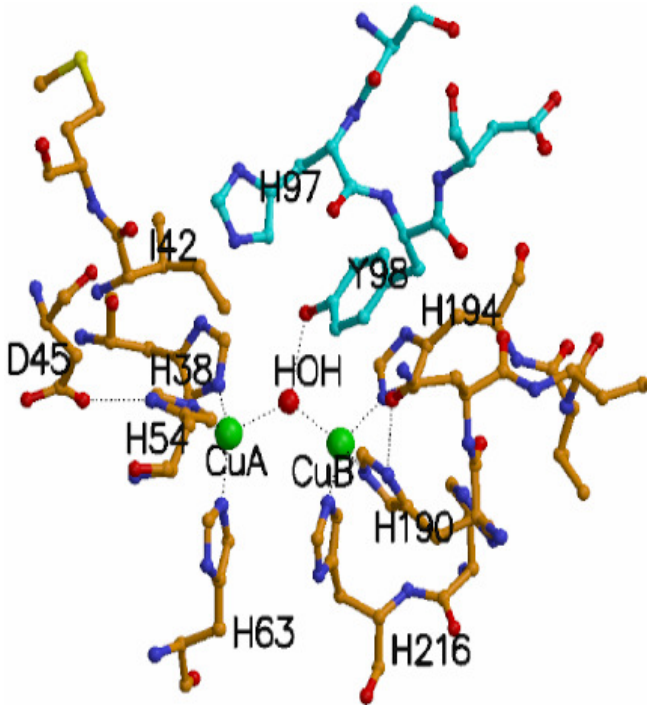


Fig. 6. Active centers of tyrosinase and of structurally homologous proteins. A, active center of the met form I of tyrosinase complexed with ORF378. Carbon atoms from the residues of tyrosinase and ORF378 are shown in orange and cyan, respectively

ACKNOWLEDGEMENTS

The collaboration of H. Parastar and Professor Jalali research group Chemometrics Lab., Department of Chemistry, Sharif University of Technology is gratefully acknowledged. The authors also appreciated the sincere collaboration of Dr. A.A. Pahlevan for improving the linguistic quality of the text.

REFERENCES

1. H.S. Mason, *Adv. Enzymol*., **16**, 105 (1955).
2. G. Prota, M. d'Ischia and A. Napolitano, in eds.: J.L. Nordlund, R. Boissy, V. Hearing, R. Kingand and J.P. Ortonne, The Pigmentary System, University Press, Oxford, p. 307 (1998).
3. K. Lerch, in ed.: H. Sigel, Metal Ions in Biological Systems, Marcel Dekker, New York, p. 146 (1981).
4. E.I. Solomon, U.M. Sundaram and T.E. Machonkin, *Chem. Rev.*, **96**, 2563 (1996).
5. J.M. Chalmers, Spectroscopy in Process Analysis, CRC, London (2000).
6. A. Nijhuis, S. de Jong and B.G.M. Vandeginste, *Chemom. Intell. Lab. Syst*., **38**, 51 (1997).
7. F. Despagne, D.L. Massart and P. Chabot, *Anal. Chem*., **72**, 1657 (2000).
8. H. Gampp, M. Maeder, C. Meyer and A.D. Zuberbuhler, *Talanta*, **33**, 943 (1986).
9. H.R. Keller and D.L. Massart, *Anal. Chim. Acta*, **246**, 379 (1991).
10. E.R. Malinowski, Factor Analysis in Chemistry, Wiley, New York, edn. 3 (2002).
11. S. Bijlsma and A.K. Smilde, *Anal. Chim. Acta*, **396**, 231 (1999).
12. R.I. Shrager, *Chemom. Intell. Lab. Syst.*, **1**, 59 (1986).
13. R. Tauler, *Chemom. Intell. Lab. Syst*., **30**, 133 (1995).
14. S. Bijlsma and A.K. Smilde, *J. Chemom*., **14**, 541 (2000).
15. H. Kamahldin, F.R. Jazii, A.A. Karkhane and S.S. Borojer, *Iranian J. Biotech.*, 2 (2004).
16. F.C. Sanchez, J. Toft, B. Van den Bogaert and D.L. Massart, *Anal. Chem*., **68**, 79 (1996).
17. W. Windig and J. Guilment, *Anal. Chem*., **63**, 1425 (1991).
18. W. Windig, C.E. Heckler, F.A. Agblevor and R.J. Evans, *Chemom. Intell. Lab. Syst*., **14**, 195 (1992).
19. R.A. Johnson and D.W. Wichern, Applied Multivariate Statistical Analysis, Englewood Cliffs, Perintice-Hall (1982).
20. G. Strang, Linear Algebra and its Applications, Harcourt Brace Jovanovich, Orlando, edn. 3 (1998).
21. H. Shen, Y. Liang, O.M. Kvalheim and R. Manne, *Chemom. Intell. Lab. Syst*., **51**, 49 (2000).
22. S. Nevea, A. de Juan and R. Tauler, *Anal. Chim. Acta,* **446**, 185 (2001).
23. M. Esteban, C. Ariño, J.M. Diaz-Cruz, M.S. Diaz-Cruz and R. Tauler, *Trends Anal. Chem.*, **19**, 49 (2000).
24. X.H. Song, A.V. Polissar and P.K. Hopke, *Atmos. Environ.*, **35**, 5277 (2001).
25. W.H. Lawton and E.A. Sylvestre, *Technometrics,* **13**, 617 (1971).
26. R. Tauler, A.K. Smilde and B.R. Kowalski, *J. Chemometr*., **9**, 31 (1995).
27. H. Shen, Y. Liang, O.M. Kvalheim and R. Manne, *Chemom. Intell. Lab. Syst*., **51**, 49 (2000).
28. A. deJuan, M. Maeder, M. Martinez and R. Tauler, *Chemom. Intell. Lab. Syst*., **54**, 123 (2000).
29. Y. Matoba, T. Kumagai, A. Yamamoto, H. Yoshitsu and M. Sugiyama, *J. Biol. Chem.*, **281**, 8981 (2006).
30. J.N. Rodriguez-Lopez, J. Tudela, R. Varon, F. Garcia-Carmona and F. Garcia-Canovas, *J. Biol. Chem.*, **267**, 3801 (1992).