Volume 252, number 1,2, 118–120 FEB 07435 July 1989

# An infrared spectroscopic study of β-galactosidase structure in aqueous solutions

J.L.R. Arrondo, A. Muga, J. Castresana, C. Bernabeu* and F.M. Goñi

*Department of Biochemistry, Faculty of Science, University of the Basque Country, PO Box 644, 48080 Bilbao* and **Centro de Investigaciones Biológicas, CSIC, Velázquez 144, 28006 Madrid, Spain*

Received 10 April 1989; revised version received 9 June 1989

Fourier-transform infrared spectroscopy has been used to elucidate the secondary structure of *E. coli* β-galactosidase in aqueous solution. The structure of this enzyme was previously unknown above the level of the amino acid sequence. Spectra have been recorded in both $H_2O$ and $D_2O$ media; mutually complementing data are obtained, that provide unambiguous structural information. The results show that β-galactosidase contains 40% β-sheet and 35% α-helical structure, with smaller proportions of random coil (12%) and β-turns (13%).

Fourier-transform infrared spectroscopy; Galactosidase, β-; Protein structure; Secondary structure; (*E. coli*)

## 1. INTRODUCTION

Fourier-transform infrared spectroscopy (FT-IR) has found extensive application in the study of protein conformations in aqueous media [1–4]. Information is usually obtained from the so-called amide I band, mostly due to peptide bond C=O stretching vibrations [5,6]. This band is constituted by the overlapping of several broad components, in turn associated with the different protein structural patterns. Band narrowing techniques (e.g. Fourier-space deconvolution and derivation) are used to decompose the amide I into its constituent bands [7]. Quantitative data of protein secondary structures have been obtained by FT-IR in $D_2O$ media that match very well the corresponding X-ray diffraction data [4]. The present paper describes the elucidation of the structure of *E. coli* β-galactosidase; despite intensive genetic and biochemical studies, the structure of this protein was not known above the sequence level, mainly because of difficulties in obtaining crystals suitable for X-ray studies [8]. IR spectra in both $H_2O$ and $D_2O$ have been examined, and the corresponding data complement each other to provide unambiguous structural information.

*Correspondence address:* J.L.R. Arrondo, Department of Biochemistry, Faculty of Science, University of the Basque Country, PO Box 644, 48080 Bilbao, Spain

## 2. MATERIALS AND METHODS

*E. coli* β-galactosidase (EC 3.2.4.7) was obtained from Boehringer-Mannheim and ran as a single band in SDS-PAGE with a molecular mass for the monomer of approx. 150 kDa. The protein was dialyzed against 10 mM ammonium bicarbonate, freeze-dried and resuspended in 10 mM Hepes buffer in either $H_2O$ or $D_2O$ with a pH or pD of 7.0. The enzyme activity (estimated using *o*-nitrophenylgalactoside as substrate and measuring changes in absorbance at 414 nm) was the same in both $H_2O$ and $D_2O$ buffers. Aqueous protein samples were measured as liquid films in a thermostatted Harrick cell (Harrick, Ossining, NY); 6 μm window spacers were used for the $H_2O$ sample and 50 μm in the case of $D_2O$. 512 and 256 scans were collected, respectively, for the $H_2O$ and $D_2O$ samples, averaged, and apodized with a Happ-Gentzel function to obtain a nominal resolution of better than 2 $cm^{-1}$. Solvent contribution was eliminated by subtracting the pure buffer spectrum from the original one. The criterion used was to maintain a straight baseline between 1300 and 1900 $cm^{-1}$ with the maximum difference factor not giving rise to negative lobes; in our case, this was equivalent to abolishing the water band at 2150 $cm^{-1}$ [9]. The spectra were then transferred to an IBM personal computer where standard Fourier self-deconvolution and Fourier derivation methods were used [10]. Band decom-



*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/89/$3.50 

position was accomplished by a least-squares method [11]. The original band contour could be exactly reconstructed by addition of the component bands, and the corresponding derivative curves were almost identical to those of the original spectrum. The criteria used to discriminate among the possible fittings that accomplished the above condition were the same as those described for triose-phosphate isomerase in $D_2O$ [12]. However, $\beta$-galactosidase being a much larger protein, the number of theoretically possible fittings was also considerably higher, therefore component band positions (as obtained from derivative spectra) were kept fixed throughout the procedure.

## 3. RESULTS AND DISCUSSION

The amide I region (1600–1700 $cm^{-1}$) of spectra in $H_2O$ and $D_2O$ media are depicted in fig.1A. In $H_2O$ a maximum at 1636 $cm^{-1}$, with a shoulder at 1655 $cm^{-1}$, is observed, whereas in $D_2O$ the maximum is located at 1634 $cm^{-1}$, and the shoulder, now at 1654 $cm^{-1}$, is decreased. This solvent effect is attributed to the shift that occurs in the random coil structure band, from 1657 $cm^{-1}$ in $H_2O$ to 1643 $cm^{-1}$ in $D_2O$ [5]. Fig.1B shows the Fourier derivatives of $\beta$-galactosidase spectra in both $H_2O$ and $D_2O$. The band positions obtained in the derivative spectra are later used for band decomposition. The collapse of the bands at 1627 and 1638 $cm^{-1}$ in $H_2O$ into a band at 1634 $cm^{-1}$ in $D_2O$ is caused by the shift of the random coil band upon deuteration. These data illustrate the need for recording spectra in both $H_2O$ and $D_2O$ in order to resolve protein structures.

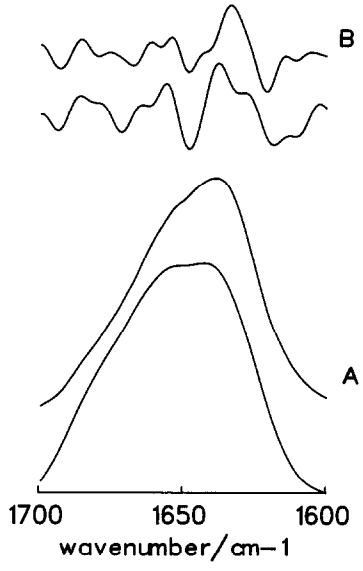


Fig.1. The 1600–1700 $cm^{-1}$ region of the infrared spectrum of $\beta$-galactosidase. (A) Spectra in $H_2O$ (lower curve) and $D_2O$ (upper curve). (B) Fourier derivatives of the $\beta$-galactosidase spectra in $H_2O$ (lower) and $D_2O$ (upper line).

Fig.2 shows the amide I band contour of $\beta$-galactosidase, in both $H_2O$ and $D_2O$, together with the band decomposition. In $H_2O$ there are protein bands at 1627 and 1638 $cm^{-1}$ ($\beta$-sheet), 1656 $cm^{-1}$ ($\alpha$-helix plus random coil), 1667 and 1678 $cm^{-1}$ ($\beta$-turns), 1686 $cm^{-1}$ ($\beta$-turns or $\beta$-antiparallel) and 1698 $cm^{-1}$ ($\beta$-edge). In $D_2O$, protein bands are located at 1634 $cm^{-1}$ ($\beta$-sheet plus random coil), 1653 $cm^{-1}$ ($\alpha$-helix), 1663 and 1685 $cm^{-1}$ ($\beta$-turns), 1673 $cm^{-1}$ ($\beta$-turns or fully exchanged $\beta$-antiparallel) and 1698 $cm^{-1}$ ($\beta$-edge not exchanged [13]). Signals around 1615 $cm^{-1}$ correspond to Tyr side chain vibrations. Bands have been assigned according to well-established criteria [1–5].

The band at 1653–1656 $cm^{-1}$ is purely due to $\alpha$-helix in $D_2O$, while containing a contribution from unordered structure in $H_2O$ [5], thus the proportion of both $\alpha$-helix and random coil structures may be derived from the experimental values in the two solvents. Signals due to $\beta$-sheet conformation appear around 1630 $cm^{-1}$ in $H_2O$ [5]. Antiparallel and parallel $\beta$-structures each show an additional signal, near 1690 and 1642 $cm^{-1}$, respectively, in $H_2O$ [5]. The signal attributed to antiparallel $\beta$-structure shifts from 1686 $cm^{-1}$ in $H_2O$ to near 1673 $cm^{-1}$ in $D_2O$, thus the presence of that conformation is marked by an increase in the relative area of the band near 1675 $cm^{-1}$ when the sample is transferred from $H_2O$ to $D_2O$ [5]. The band at

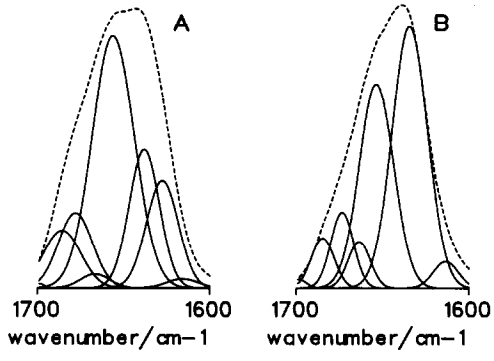


Fig.2. The amide I band contour (broken lines) and component bands (solid lines) of the infrared spectrum of $\beta$-galactosidase. (A) Spectrum in $H_2O$. (B) Spectrum in $D_2O$.

Table 1

Frequencies ($\nu$), bandwidths at half-height ($\Delta\nu_{1/2}$) and fractional band areas ($A$) of amide I band components of $\beta$-galactosidase in $H_2O$ and $D_2O$

| $H_2O$ | | | $D_2O$ | | |
|---|---|---|---|---|---|
| $\nu$ ($cm^{-1}$) | $\Delta\nu_{1/2}$ ($cm^{-1}$) | $A$ (%) | $\nu$ ($cm^{-1}$) | $\Delta\nu_{1/2}$ ($cm^{-1}$) | $A$ (%) |
| 1616 | 21 | | 1615 | 17 | |
| 1627 | 22 | 15 | 1634 | 24 | 45 |
| 1638 | 20 | 18 | 1653 | 23 | 35 |
| 1656 | 29 | 46 | 1663 | 15 | 5 |
| 1667 | 20 | 2 | 1673 | 15 | 9 |
| 1678 | 23 | 11 | 1685 | 16 | 6 |
| 1686 | 25 | 9 | 1698 | 10 | <1 |
| 1698 | 6 | <1 | | | |

The values of A are rounded off to the nearest integer. The curve-fitting procedure provides results with an estimated error of less than ± 5% of the given values, when different parameters are used at the start of the iterative process

1698 $cm^{-1}$, almost unaffected by solvent exchange, has been related to $\beta$-edge structure, i.e. $\beta$-sheet residues that are not hydrogen bonded to another $\beta$-strand but to amino acid side chains [12,13]. Most of the $\beta$-structure in this protein must be in the parallel conformation, because of the stronger intensity of the 1638 $cm^{-1}$ band (in $H_2O$) with respect to the higher frequency component of the $\beta$-antiparallel signal, at 1686 $cm^{-1}$. However, the proportion of each of these two structures is difficult to ascertain, because their molar extinction coefficients are not known. With the above considerations in mind, a structure for $\beta$-galactosidase in aqueous solution is derived from our experimental data, as shown in table 2. At present, FT-IR data cannot be directly analyzed to solve tertiary structures; however, the joint presence of $\beta$-parallel, $\beta$-edge and $\alpha$-helix features may suggest that an $\alpha/\beta$ barrel could be a domain of $\beta$-galactosidase [12]. The present work provides an example of FT-IR as a method for the quantitative description of secondary structures of proteins in solution.

Table 2

The secondary structure of *E. coli* $\beta$-galactosidase as derived from FT-IR measurements

| Structure | Distribution (%) |
|---|---|
| $\alpha$-Helix | 35 |
| $\beta$-Sheet | 40 |
| Random | 12 |
| $\beta$-Turns | 13 |

% values have been rounded off to the nearest integer

*Acknowledgements:* This work was supported in part by CAICYT (grant no.992/84) and the University of the Basque Country (grant no.042.310-0068/88).

## REFERENCES

[1] Lee, D.C. and Chapman, D. (1986) Biosci. Rep. 6, 235–255.
[2] Susi, H. and Byler, D.M. (1986) Methods Enzymol. 130, 290–311.
[3] Surewicz, W.C. and Mantsch, H.H. (1988) Biochim. Biophys. Acta 952, 115–130.
[4] Byler, M. and Susi, H. (1986) Biopolymers 25, 469–487.
[5] Susi, H., Timasheff, S.N. and Stevens, L. (1967) J. Biol. Chem. 242, 5460–5466.
[6] Parker, F.S. (1983) Applications of Infrared, Raman and Resonance Raman Spectroscopy in Biochemistry, Plenum, New York.
[7] Mantsch, H.H., Casal, H.L. and Jones, R.N. (1986) in: Spectroscopy of Biological Systems (Clark, R.J.H. and Hester, R.E. eds) pp.1–46, Wiley, New York.
[8] Littke, W. and John, C. (1984) Science 225, 203–204.
[9] Powell, J.R., Wasacz, F.M. and Jakobsen, R.J. (1986) Appl. Spectrosc. 40, 339–344.
[10] Moffat, D.J., Kauppinen, J.K., Cameron, D.G., Mantsch, H.H. and Jones, R.N. (1986) Computer Programs for Infrared Spectroscopy, NRCC Bulletin no.18, Ottawa, Canada.
[11] Fraser, R.D.B. and Suzuki, E. (1966) Anal. Chem. 38, 1770–1773.
[12] Castresana, J., Muga, A. and Arrondo, J.L.R. (1988) Biochem. Biophys. Res. Commun. 152, 69–75.
[13] Arrondo, J.L.R., Young, M. and Mantsch, H.H. (1988) Biochim. Biophys. Acta 952, 261–268.