Fourier transform infrared spectroscopy analysis of the conformational quality of recombinant proteins within inclusion bodies

ARTICLE *in* BIOTECHNOLOGY JOURNAL · FEBRUARY 2008

Impact Factor: 3.49 · DOI: 10.1002/biot.200700238 · Source: PubMed

5 AUTHORS, INCLUDING:

Diletta Ami
Università degli Studi di Milano-Bicocca
34 PUBLICATIONS **1,154** CITATIONS
SEE PROFILE

Antonino Natalello
Università degli Studi di Milano-Bicocca
68 PUBLICATIONS **1,140** CITATIONS
SEE PROFILE

Pietro Gatti-Lafranconi
Illumina, United Kingdom
19 PUBLICATIONS **391** CITATIONS
SEE PROFILE

Marina Lotti
Università degli Studi di Milano-Bicocca
83 PUBLICATIONS **1,859** CITATIONS
SEE PROFILE

Review

Fourier transform infrared spectroscopy analysis of the conformational quality of recombinant proteins within inclusion bodies

Silvia Maria Doglia, Diletta Ami, Antonino Natalello, Pietro Gatti-Lafranconi and Marina Lotti*

Dipartimento di Biotecnologie e Bioscienze, Università degli Studi di Milano-Bicocca, Milano, Italia

The solubility of recombinant proteins produced in bacterial cells is considered a key issue in biotechnology as most overexpressed polypeptides undergo aggregation in inclusion bodies, from which they have to be recovered by solubilization and refolding procedures. Physiological and molecular strategies have been implemented to revert or at least to control aggregation but they often meet only partial success and have to be optimized case by case. Recent studies have shown that proteins embedded in inclusion bodies may retain residual structure and biological function and question the former axiom that solubility and activity are necessarily coupled. This allows for a switch in the goals from obtaining soluble products to controlling the conformational quality of aggregated proteins. Central to this approach is the availability of analytical methods to monitor protein structure within inclusion bodies. We describe here the use of Fourier transform infrared spectroscopy for the structural analysis of inclusion bodies both purified from cells and *in vivo*. Examples are reported concerning the study of kinetics of aggregation and structure of aggregates as a function of expression levels, temperature and co-expression of chaperones.

Received 15 November 2007
Revised 14 December 2007
Accepted 17 December 2007

Keywords: Fourier transform infrared spectroscopy · Inclusion bodies · Recombinant proteins · Solubility

1 Biological and biotechnological aspects of the aggregation of recombinant proteins expressed in bacteria

Despite the development of a number of cell factories for the production of recombinant proteins, the Gram-negative bacterium *Escherichia coli* still holds as one of the favorite hosts because of its ability to grow rapidly and at high density on inexpensive substrates with high production yields. In the experimental and production praxis, unless the recombinant product requires specific post-translational modifications not performed by prokaryotic cells, such advantages are assumed to counterbalance the main limitation related to the use of bacterial hosts, the deposition of overproduced proteins in intracellular or periplasmatic insoluble aggregates, named inclusion bodies (IBs) [1, 2]. As a matter of fact, most recombinant proteins undergo aggregation upon expression in *E. coli*. The physiology of intracellular aggregation is known in its general and molecular aspects. Indeed, non-native protein conformations exposing hydrophobic patches prone to aggregation are commonly produced even during the *de novo* synthesis of proteins and their intracellular transport. A complex folding control machinery, composed by chaperones acting in the different steps of folding with different mechanisms of action, is in charge of preventing

Correspondence: Marina Lotti, Dipartimento di Biotecnologie e Bioscienze, Università degli Studi di Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italia
E-mail: marina.lotti@unimib.it
Fax: +39-02-64483565

Abbreviations: **ATR**, attenuated total reflection; **IB**, inclusion body; **FT-IR**, Fourier transform infrared spectroscopy

**Current address*: Consortium for Genomic Technologies (COGENTECH), Biochemistry Unit, Via Adamello 16, 20139 Milano, Italy

aggregation. A recent classification distinguishes "folding chaperones" assisting the folding of newly synthesized and misfolded proteins (*i.e.*, DnaK/DnaJ and GroEL/GroES) and "holding chaperones" that bind to partially folded proteins thereby preventing overloading of the chaperones systems (*i.e.*, IbpA and IbpB). Other chaperones are active in the disaggregation of IBs and in the rescue of proteins by active solubilization of aggregates [3]. Stress conditions might overload cells with misfolded or partly folded proteins exceeding the buffering capacity of the folding control machinery and result in protein aggregation. In the cell context, the overexpression of recombinant proteins induces a severe metabolic and physiological stress [4], leading to a temporary crash in the folding ability of the cells. Consistent with this view is the idea that IBs would be a strategy for protecting cells from the exposition to toxic misfolded or partly folded proteins. Accordingly, it has been shown that aggregates are dissolved when the stress situation is relieved [5].

The observation that IBs mainly consist of a single protein species suggested that aggregation is likely to arise from selective intermolecular interactions through homologous protein patches [6]. From a practical point of view, this identifies IBs as reservoirs of relatively pure proteins sequestered in proteolysis-resistant aggregates from which they can be recovered after appropriate processing [1, 2]. For this reason, downstream, aggregation can be considered as an advantage for production, provided the native protein can be easily refolded after solubilization of IBs purified from cells [2, 7]. However, such procedures might be cumbersome and impact on the costs and yields of the process, especially on the industrial scale [8].

Aggregation is difficult to prevent and even to predict on the basis of the amino acid sequence or other protein features, although a few systematic studies have appeared in recent years [9]. Thus, strategies for controlling protein solubility strongly depend on the specific protein and have to be developed case by case with a time-consuming trial and error approach [1, 10]. For example, intracellular aggregation can be partly controlled by decreasing the rate of heterologous protein synthesis, *i.e.*, at low production temperatures, by the use of weak promoters or, whenever inducible promoters drive expression, by sub-optimal amounts of inducers. Such physiology-based approaches might allow a shift in the ratio of insoluble to soluble protein [10]. Further improvements have been achieved by fusing the protein of interest with a partner endowed with high solubility or by the co-expression of molecular chaperones [10]. The former strategy relies on the observation that the solubility of a protein might be strongly enhanced by a solubilizing fusion partner such as maltose binding protein, N-utilizing substance A, glutathione S-transferase (GST) and thioredoxin, but requires an accurate selection of the partner protein [11] and proteolytic cleavage of the fusion product followed by the purification of the protein of interest [12]. Co-expression of folding modulators with the recombinant protein has provided interesting improvements. Unfortunately, neither a single chaperone nor a defined combination of such proteins is able to act as a "universal foldase" and the success of co-expression is strongly dependent on interactions between the target protein and one or more groups of specific co-operative folding helpers acting in a highly coordinated way. It is therefore necessary to identify productive protein/chaperones combinations and to tune the timing and level of co-expression, with a trial and error process avoiding possible undesirable effects, such as toxicity of the overexpressed chaperones or competition of the two overexpressed proteins for the synthetic resources of the host cell [13, 14]. Mutagenesis has been also exploited as a tool to control aggregation, as the specific amino acid sequence affects protein stability, solubility and, as a consequence, propensity to aggregation [15–17]. Several bioinformatic tools have been developed to predict aggregation "hot spots" in amino acid sequences and support mutant design [16, 18–20]. Beyond intrinsic properties of the polypeptide, other factors play a role in its aggregation and are related to the molecular and physiological cellular environment where folding (or aggregation) takes place.

To date, important improvements have been achieved with one or a combination of the methods mentioned above, but a complete reversion of aggregation to solubility has seldom been obtained. As a consequence, the recombinant product is often partitioned between aggregates and soluble proteins, thus reducing the amount of active polypeptides.

2 Proteins within IBs may retain residual structure and function

In recent years, the picture of the molecular architecture and physiological role of IBs has considerably changed. Seminal studies have shown that IBs, far from being amorphous deposits of denatured proteins as they were described until a few years ago, are indeed heterogeneous and dynamic structures [21, 22]. Aggregation is no longer considered

as a dead end for proteins since it might be reversed by disaggregating chaperones when the stress is relieved, *i.e.*, when the rate of protein synthesis decreases [22]. In this context, an active role of DnaK in the unfolding and disassembly of proteins within stable aggregates has been clearly demonstrated [23–24]. Even more interesting was the finding that the structure of aggregates varies according to the protein and, for the same protein, to the physiological conditions applied during its expression. Proteins entrapped in aggregates are not necessarily totally denatured but they can retain residual secondary structure and biological activity [25–29]. Such studies provide important advances in the understanding of the mechanisms of aggregation and have revolutionized the current view on the biological significance and structure of aggregated proteins, opening the way for knowledge-based improvements of biotechnological process. This new information, joined to the finding that proteins may indeed form inactive, although soluble, aggregates [10, 30] led to the development of a new concept modifying the simplified classification of recombinant proteins being soluble (therefore active) or insoluble (therefore inactive). This refers to the "quality" of the recombinant protein, which, beyond solubility also considers its actual native conformation and biological activity [31–34]. This conceptual scenario implies that it should be possible to manipulate aggregated proteins to uncouple solubility from the retaining of structure, and thus from activity [34]. In other words, the goal would switch from obtaining soluble polypeptides to improving the conformational "quality" of proteins within IBs. Physiological and genetic methods described above might fail in reversing aggregation, but eventually be of advantage to increase the residual secondary structure and activity of aggregated proteins. The ability to produce IBs containing native-like proteins opens interesting scenarios, for example, it would allow the use of milder solubilization protocols [35, 36] during the refolding but would also foster the development of novel biocatalytic strategies. It is possible to envisage biotranformations in which native-like, activity-retaining enzyme aggregates are simply recovered after cell disruption by centrifugation and/or filtration and directly used as biocatalysts [29]. Should the hypothesis hold true that aggregated proteins are resistant not only to proteases but also to exposure to severe experimental conditions, this method might turn out to be particularly favorable in processes at high temperatures or at high concentrations of organic solvents [37, 38].

This recent focus on the importance of the structure of proteins embedded in aggregates has prompted the implementation of analytical methods to study the kinetics of the formation and the fine architecture of IBs. Several biophysical techniques are useful, *e.g.*, the binding to fluorescent dyes specific for structured aggregates (*e.g.*, thioflavin-T), circular dichroism (CD), transmission and scanning electron microscopy, and mass spectrometry [39, 40]. In this context, Fourier transform infrared spectroscopy (FT-IR) is particularly powerful. FT-IR has been broadly applied for the study of thermal stability and aggregation of proteins in solution [41–43] and has been recently implemented for the analysis of aggregated proteins within intact cells [26, 44, 45]. This review article highlights the potentiality of FT-IR in the study of key aspects of the aggregation of recombinant proteins, *i.e.*, (i) monitoring the kinetics of IBs formation within living cells, (ii) evaluating the secondary structure of aggregated proteins directly inside cells, and (iii) inspecting the fine structure of isolated IBs. The possibility for analyzing samples taken directly from the culture without preliminary processing, allows the dynamic monitoring of the process of aggregation *in vivo*, suggesting further implementations towards the on-line application of this technique.

3 FT-IR spectroscopy for the study of protein aggregation

3.1 Method

Over the last decade the potential of FT-IR spectroscopy for protein science has been widely recognized in investigating protein secondary structure, folding and stability [41, 43, 46, 47]. Furthermore, protein aggregation, a central issue in medical sciences [48] and biotechnology [33], has been successfully approached by this technique [47, 49, 50]. In particular, protein secondary structure and aggregation can be studied using the infrared spectrum of protein in the amide I region (1700–1600 cm^{-1}), which is due to the C=O peptide bond absorption, since a specific (FT-IR) infrared response characterizes protein–protein interaction in aggregates [49, 50]. Indeed, not only proteins in solution but also insoluble aggregates and other highly scattering systems (such as intact cells [51–53]) can be studied, since FT-IR spectroscopy is not affected by scattering disturbances. Previous studies by ourselves and others have shown that the specific infrared marker band of aggregation can also be identified in intact cells [26, 44, 45] and tissues [54–56]. Therefore, both the kinetics of aggregation and the quality of aggregated proteins

can be followed upon changing different parameters. Here we describe, with some experimental detail, how to apply FT-IR to the analysis of aggregates both within intact cells and after extraction, using as an example *E. coli* cells either as wild type or producing recombinant interferon (IFN)-α2b at different levels [44]. Changes in the relevant spectra are therefore related to the level of expression of the recombinant protein. FT-IR analysis can be directly performed on small aliquots of the culture broth (1–2 mL) centrifuged to isolate the cell pellet and then resuspended in 20–1000 μL distilled water. Small volumes (~5 μL) are deposited on a BaF_2 infrared support. After drying at room temperature for 20–30 min to eliminate bulk water, the cell spectrum is collected in transmission by selecting a sample area of about 100 × 100 μm through the variable microscope aperture. The infrared absorption spectrum shown in Fig. 1 was produced by intact transformed *E. coli* cells and emphasizes the kind of information can be obtained by this technique. In the experiments summarized here, a shoulder of the main amide I absorption band was detected around 1628 cm^{-1} in the spectrum of the high producer strain (Fig. 1a). This band component was better appreciated after a second derivative analysis of the spectrum (Fig. 1b). This mathematical procedure enables the broad amide I band to be resolved into its overlapping components caused by the different secondary structures and aggregates of the total cell protein content. The peak positions and intensities [42] of these components can be taken from the negative bands of the second derivative spectrum. The second derivative spectrum of the control strain displayed two major components at 1658 and 1638 cm^{-1} that can be, respectively, assigned to α-helix and β-sheet structures of the total cell proteins. In the higher producer strain, a new band around 1628 cm^{-1} (that can be assigned to the intermolecular β-sheet protein–protein interaction) was observed. This marker band of aggregation grows with the level of expression of the recombinant protein deposited in the form of IBs [44]. Therefore, the kinetics of protein aggregation can be monitored within intact cells, in a rapid and simple way, under different culture conditions, as illustrated in the following. The statistical significance of the spectral data was assessed by cluster analysis performed on the second derivative spectra of the wild-type and producer strains, taken from three independent cellular preparations. The Ward's algorithm and Euclidean distance measure were applied, using the Statistics Tool box of Matlab 5.3 (The Math Works, Inc., Natick, MA). The dendrogram, obtained in the wide range from 1800 to 900 cm^{-1}, showed a clear separation into three distinct clusters for the spectra obtained from wild-type and recombinant *E. coli* cells (Fig. 1c). A chemometric analysis of FT-IR absorption spectra of *E. coli* cells overexpressing a recombinant protein has been also recently performed by principal component analysis [57]. When examining low producer strains (Figs. 1a and b), the marker band of aggregation was not resolved even by the second derivative analysis of the spectrum. Under conditions of low protein expression an additional analysis is required. To determine the response of the recombinant protein it is necessary to reduce the contribution of the other cell proteins, by subtracting the control strain spectrum from that of the producer strain [26, 44, 45].

In addition to the analysis of aggregation in intact cells, FT-IR spectroscopy of extracted IBs provided important structural information about the recombinant protein within the aggregates. After extraction and rinsing, IBs were deposited on an infrared transparent window and, after water evaporation, were measured in transmission as described [58]. The infrared response of IBs in the amide I region allowed the protein–protein interaction to be characterized within aggregates through the peak position and relative intensity of the infrared band around 1630–1625 cm^{-1}, which reflect the compactness of IBs and the extent of aggregation, respectively [26]. Moreover, secondary structure components, with peak positions close to those of the native protein, were observed in the spectrum, indicating that the protein was not completely unfolded but rather retained a native-like structure within IBs. As these structural features correlate with functional properties, as for example biological activity and sensitivity to mild solubilization procedures during refolding [32, 36, 59], the relevance of an easy method to assess them is of obvious interest. We should add that similar investigations can be performed by working in attenuated total reflection (ATR), a sampling technique extensively used in infrared spectroscopy [28, 32, 34, 60].

3.2 Effects of temperature on kinetics of aggregation and protein structure within IBs

The temperature during overexpression can affect the protein partitioning between soluble and insoluble fractions, and can be accordingly adjusted to maximize the level of the protein expressed in a native soluble form. As the FT-IR method allows monitoring the kinetics of protein aggregation within intact cells, the effect of temperature on IBs formation can be also studied. We illustrate here this approach through results obtained with pro-

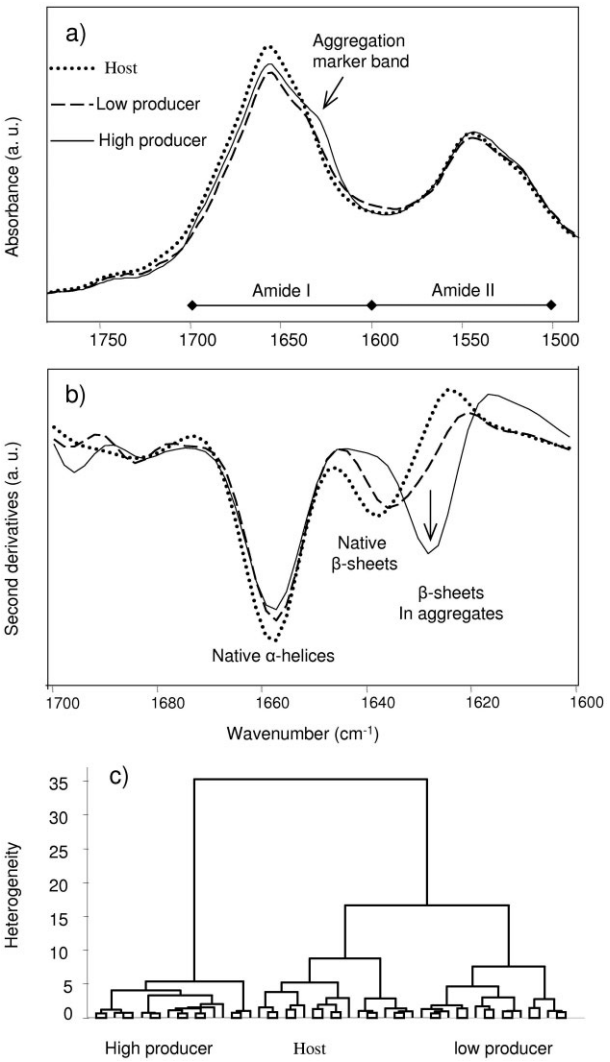

Figure 1. FT-IR marker band of aggregation in intact cells. (a) FT-IR absorption spectra in the amide I and amide II regions of the host, the low producer and the high producer *E. coli* strains. The arrow indicates the shoulder due to the protein aggregate absorption. (b) Second derivatives in the amide I region of the absorption spectra reported in (a), normalized at the ~1517 cm^{-1} tyrosine band; native-like secondary structures and aggregation marker band are indicated. (c) Cluster analysis of the second derivative spectra of the host, the low producer and the high producer *E. coli* strains.

teins with different intrinsic properties. The lipase from the psychrotrophic bacterium *Pseudomonas fragi* (PFL) is a cold-adapted protein, characterized by high instability at moderate temperature and is therefore a sensitive tool to investigate temperature-related effects. The second protein is a fusion between green fluorescent protein (GFP) and GST. GST is often employed as a solubility partner; on the other hand, GFP fluorescence is a reporter of its native or native-like structure. PFL was expressed at 27°C and 37°C (a temperature that is strongly destabilizing for the protein) and the spectra recorded at different times from the induction of expression were compared to those of non-transformed control strains [45]. Figure 2a shows the second derivatives of the spectrum obtained with the producer strain at 37°C as resulting after normalization at the 1517 cm^{-1} tyrosine band to compensate for possible difference in protein content in different samples. A shoulder around 1627 cm^{-1}, due to protein aggregates, was observed to increase with the time elapsed from induction. As discussed above, this component can be better resolved by subtracting the spectrum of the control cells from that of the producer cells. The second derivatives of the subtracted spectra taken at three selected times after induction at 37°C and 27 °C are reported in Figs. 2b and c, where the marker band of aggregation occurred around 1627 cm^{-1}. This aggregate component increased with time so that the kinetics of IB formation could be monitored, and these reached higher values for the culture grown at 37°C at any considered time (Fig. 2d). A higher rate of aggregation was observed at 37°C with similar time dependence when data were analyzed by the subtractive procedure (Fig. 2b) or only by the simplified method (Fig. 2a). SDS-PAGE analysis of cultures induced at 37°C, 27°C or 17°C confirmed the data obtained by FT-IR showing that, while IBs were clearly present at all tested temperatures their relative amount was considerably temperature dependent and a band corresponding to PFL was visible in the soluble fraction only when cells are grown at lower temperature (Fig. 2e). The influence of temperature on the conformational properties of the protein entrapped within IBs was studied on extracted aggregates. As shown by the second derivative spectra in Fig. 3a, aggregated PFL retained a higher content of native-like secondary structure when produced at 27°C, while aggregation increased with temperature. The same IBs were tested for residual activity in a standard assay for lipase activity. In good agreement with the hints provided by the spectroscopic analysis, IBs obtained at low temperature (17°C) were by fare more active than those extracted at 37°C (not shown).

To investigate the temperature dependence of the conformational quality of aggregates in more depth, similar experiments were performed on the fusion GFP-GST [30] that is less temperature sensitive and more soluble than PFL. In this case, the recombinant protein was produced at 20°C and 30°C. SDS-PAGE analysis on soluble proteins and extracted IBs (Fig. 3c) indicated that the overall production at the two temperatures was similar but the partitioning between soluble and insoluble

Biotechnol. J. 2008, 3, 193–201

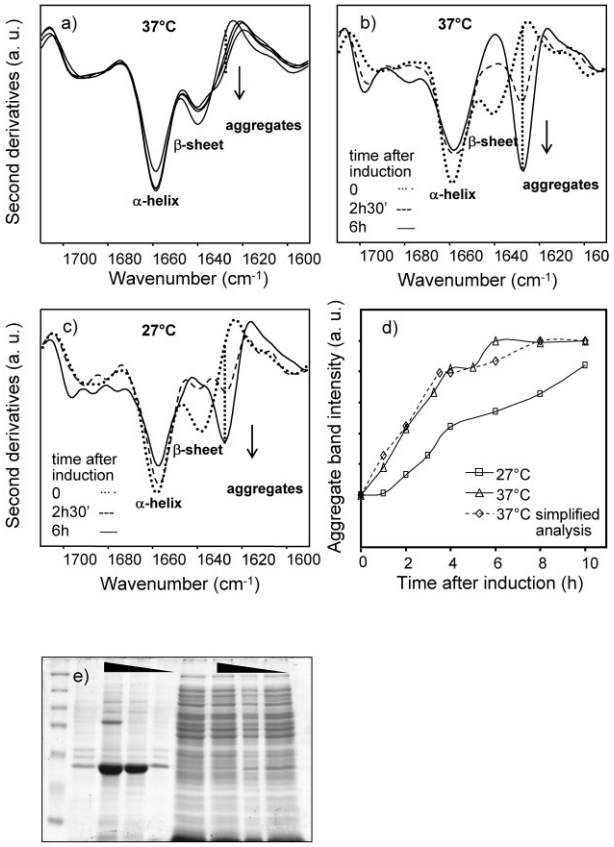

Figure 2. Kinetics of aggregation in intact cells by FT-IR spectroscopy. (a) Second derivative FT-IR spectra of *E. coli* expressing PFL at 37°C taken at different times after induction. The intensity of the shoulder due to protein aggregation (dotted line) is reported *versus* time in (d). (b) Second derivative FT-IR spectra of cells expressing PFL at 37°C after subtraction of absorption spectra of control cell. Selected spectra taken at different times after induction are reported. (c) Second derivative FT-IR spectra cells producing PFL at 27°C after subtraction of control cell spectra. (d) Kinetics of IB formation at 27°C and 37°C. The intensity of the marker band of aggregation, indicated by the dotted line in (a–c), is reported *versus* time after induction. Arrows point to increasing time after induction and all spectra were normalized at the ~1517 cm^{-1} tyrosine band. (e) SDS-PAGE analysis of IBs (lanes 2–5) and soluble protein fractions (lanes 6–9) from cells producing recombinant PFL either not induced (first lane of each set) or induced at 37, 27, or 17°C. Lowering of the temperature applied during fermentation is marked by the direction of the arrow. First lane: molecular marker band.

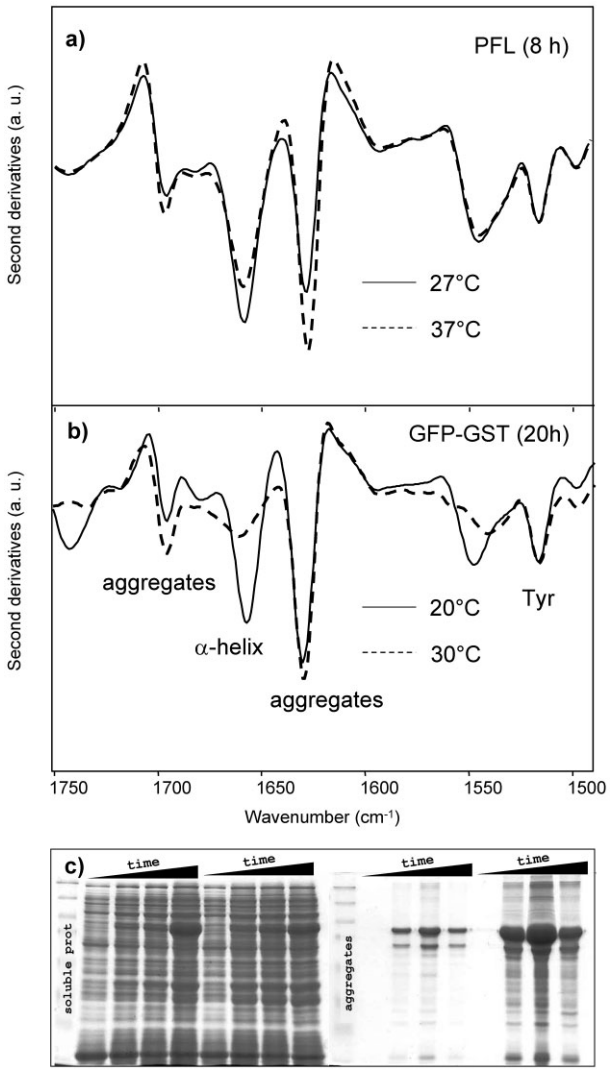

Figure 3. Structural properties of extracted IBs. (a) Second derivative FT-IR spectra of PFL-IBs extracted 8 h after induction from the 27 and 37°C fermentations. (b) Second derivative FT-IR spectra of GFP-GST IBs extracted 20 h after induction from fermentations at 20 and 30°C. All spectra were normalized at the ~1517 cm^{-1} tyrosine band. (c) SDS-PAGE of soluble and insoluble fractions from *E. coli* producing the GFP-GST fusion at 20°C (dark gray arrow) and 30°C (gray arrow). Samples are taken before induction (first lane of each set) and at times growing from left to right.

fraction was shifted towards aggregates at 30°C. FT-IR performed on purified IBs pointed to an extensive loss of residual secondary structure in the 30°C aggregates, that could be appreciated in particular from the decrease of the α-helical component (Fig. 3b). Accordingly, GFP fluorescence detected by confocal microscopy was reduced in the sample fermented at 30°C (not shown). Thus, the biochemical evidence of a shift in the ratio between soluble and insoluble proteins induced by temperature was well complemented by the spectroscopic data showing a different conformational quality of the aggregates; this information in turn related to the activity of the aggregated proteins. Similar results have recently been reported for the GFP expressed alone in *E. coli* at 16, 30, 37°C. The infrared data indicated that an appreciable native-like structure was only present in the protein within IBs at low temperature. The aggregation band in this case also increased and shifted toward lower wavenumbers as the temperature increased, indicating a higher compactness of the aggregates [28].

3.3 Effect of chaperones on the protein structural properties within IBs

Chaperones and proteases are known to play an important role in protein folding and in the quality control cell machinery. Their deletion or overexpression in recombinant bacteria is, therefore, expected to affect the fractioning of protein into soluble and insoluble form, with possible implications in the protein structural properties within IBs. Unexpectedly, the specific fluorescence of an engineered GFP was found to increase when chaperones and proteases were deleted in recombinant *E. coli* cells [34], indicating the presence of protein in its functional form. In addition to fluorescence measurements, the authors employed FT-IR/ATR spectroscopy to characterize the protein structural properties within IBs produced in wild-type *E. coli* cells, and in a number of recombinant strains deficient in DnaK, GroEL, GroES, ClpB, ClpP and Lon proteins. A down shift of the aggregated band (from 1627 to 1623 cm^{-1}) was observed, showing that IBs produced in the chaperones-deficient strains were more compact that IBs deposited in wild-type cells. In spite of high aggregation, the fluorescence associated to IBs was found to be very high in absence of DnaK, ClpA, ClpB, and ClpP. Accordingly, the FT-IR spectra detected the presence of native-like secondary components [34]. These properties were also revealed by FT-IR/ATR spectroscopy for a recombinant β-galactosidase expressed in the absence of DnaK [32], suggesting that they are not peculiar to the GFP model system.

4 Conclusive remarks

The study of IB formation, accumulation and intrinsic properties is currently extremely popular, given the strong connections of this process with protein folding, solubility, activity and its application in the biotechnological industry. There still is, however, a lack in knowledge that need to be filled before *in vivo* aggregation can be fully understood and exploited. Biochemical, physiology-based and biophysical approaches have been developed to address different aspects of the topic and it is now clear that there is no optimal method to comprehensively investigate IBs, but rather a set of available and powerful tools. Among them, FT-IR combines a method to study the structural properties of aggregates with one for fast and easy monitoring of the kinetics of aggregation, thus providing the possibility of controlling the production process to obtain a high conformational quality of proteins within aggregates.

This work was supported by grants F.A.R. (Fondo di Ateneo per la Ricerca) of the University of Milano-Bicocca to S.M.D. and M.L. A.N. acknowledges a fellowship by Progetto Ingenio granted by Fondo Sociale Europeo, Ministero del Lavoro e della Previdenza Sociale and Regione Lombardia and P. G.-L. a fellowship by the Fondazione Fratelli Confalonieri. The authors gratefully thank A. De Marco for providing the GFP-GST construct and for helpful discussions.

The authors have declared no conflict of interest.

5 References

[1] Baneyx, F., Mujacic, M., Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* 2004, *22*, 1399–1408.

[2] Fahnert, B., Lilie, H., Neubauer, P., Inclusion bodies: Formation and utilisation. *Adv. Biochem. Eng. Biotechnol.* 2004, *89*, 93–142.

[3] Schlieker, C., Bukau, B., Mogk, A., Prevention and reversion of protein aggregation by molecular chaperones in the *E. coli* cytosol: Implications for their applicability in biotechnology. *J. Biotechnol.* 2002, *96*, 13–21.

[4] Hoffmann, F., Rinas, U., Stress induced by recombinant protein production in *Escherichia coli*. *Adv. Biochem. Eng. Biotechnol* . 2004, *89*, 73–92.

[5] Villaverde, A., Carrió, M. M., Protein aggregation in recombinant bacteria: Biological role of inclusion bodies. *Biotechnol. Lett.* 2003, *25*, 1385–1395.

[6] Speed, M. A., Wang D. I., King J., Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion bodies composition. *Nat. Biotechnol.* 1996, *14*, 1283–1287.

[7] Vallejo, L. F., Rinas, U., Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb. Cell Fact*. 2004, *3*, 11.

[8] Clark, E. D., Protein refolding for industrial processes. *Curr. Opin. Biotechnol*. 2001, *12*, 202–207.

[9] Koschorreck, M., Fischer, M., Barth, S., Pleiss, J., How to find soluble proteins: A comprehensive analysis of alpha/beta hydrolases for recombinant expression in *E. coli*. *BCM Genomics* 2005, *6*, 49.

[10] Sørensen, H. P., Mortensen, K. K., Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb. Cell. Fact*. 2005, *4*, 1.

[11] Dummler, A., Lawrence, A. M., de Marco, A., Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb. Cell Fact.* 2005, *4*, 34.

[12] Waugh, D. S., Making the most of affinity tags. *Trends Biotechnol.* 2005, *23*, 316–320.

[13] De Marco, A., De Marco, V., Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning. *J. Biotechnol.* 2004, *109*, 45–52.

[14] De Marco, A., Vigh, L., Diamant, S., Goloubinoff, P., Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolites, plasmid- or benzyl alcohol-over-

Marina Lotti is the head of the laboratory of Protein Engineering and Enzymology at the Department of Biotechnology and Biosciences of the University of Milano-Bicocca. She graduated in Biology in Milano and received her PhD in Natural Sciences at the Max-Planck-Institute of Berlin. Presently she is professor of Biochemistry at the University of Milano-Bicocca. Major research interests concern the relationships among sequence, conformation, specificity and stability of proteins. These studies are based on a combined approach of mutagenesis, bioinformatics, biochemical assays and are well complemented by the biophysical study of protein conformation carried out in collaboration with other research groups.

Silvia Maria Doglia is Associate Professor of Physics at the Faculty of Sciences of the University of Milano-Bicocca. At the Department of Biotechnology and Biosciences, her research activity in Biophysics focuses on the study of structure, stability and aggregation of proteins by optical spectroscopies and the in vivo study by Fourier transform infrared spectroscopy. In the past years she carried out nucleic acid research as Visiting Professor at the Universities of Orléans and Reims in France, and as Visiting Scientist at the University of Stockholm in Sweden. A recent research project addressed the role of mt DNA in tumor cells by confocal fluorescence microscopy.

expressed molecular chaperones. *Cell Stress Chaperones* 2005, *10*, 329–339.

[15] Pace, C. N., Trevino, S., Prabhakaran, E., Scholtz, J. M., Protein structure, stability and solubility in water and other solvents. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. 2004, *359*, 1225–1234.

[16] Ventura, S., Sequence determinants of protein aggregation: Tools to increase protein solubility. *Microb. Cell Fact*. 2005, *4*, 11.

[17] Srisailam, S., Kumar, T. K., Srimathi, T., Yu, C., Influence of backbone conformation on protein aggregation. *J. Am. Chem. Soc.* 2002, *124*, 1884–1888.

[18] Fernandez-Escamilla, A., M., Rousseau, F., Schymkowitz, J., Serrano, L., Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* 2004, *22*, 1302–1306.

[19] Linding, R., Schymkowitz, J., Rousseau, F., Diella, F., Serrano, L., A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. *J. Mol. Biol.* 2004, *342*, 345–353.

[20] Chiti, F., Stefani, M., Taddei, N., Ramponi, G., Dobson, C. M., Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 2003, *424*, 805–808.

[21] Carrió, M. M., Cubarsi, R., Villaverde, A., Fine architecture of bacterial inclusion bodies. *FEBS Lett*. 2000, *471*, 7–11.

[22] Carrió, M. M., Villaverde, A., Construction and deconstruction of bacterial inclusion bodies. *J. Biotechnol.* 2002, *96*, 3–12.

[23] Ben-Zvi, A., De Los Rios, P., Dietler, G., Goloubinoff, P., Active solubilization and refolding of stable protein aggregates by cooperative unfolding action of individual Hsp70 chaperones. *J. Biol. Chem.* 2004, *279*, 37298–37303.

[24] De Los Rios, P., Ben-Zvi, A., Slutsky, O., Azem, A., Goloubinoff, P., Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proc. Natl. Acad. Sci. USA* 2006, *103*, 6166–6171.

[25] Oberg, K., Chrunyk, B. A., Wetzel, R., Fink. A. L., Native-like secondary structure in interleukin-1 beta inclusion bodies by attenuated total reflectance FTIR. *Biochemistry* 1994, *33*, 2628–2634.

[26] Ami, D., Natalello, A., Taylor, G., Tonon, G., Doglia, S. M., Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy. *Biochim. Biophys. Acta* 2006, *1764*, 793–799.

[27] Tokatlidis, K., Dhurjati, P., Millet, J., Beguin, P., Aubert, J. P., High activity of inclusion bodies formed in *Escherichia coli* overproducing *Clostridium thermocellum* endoglucanase D. *FEBS Lett.* 1991, 282, 205–208.

[28] Vera, A., Gonzáles-Montalbán, N., Arís, A., Villaverde, A., The conformational quality of insoluble recombinant proteins is enhanced at low growth temperature. *Biotechnol. Bioeng.* 2007, *96*, 1101–1106.

[29] Garcia-Fruitos, E., Gonzàlez-Montalbàn, N., Morell, M., Vera, A. *et al*., Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb. Cell Fact*. 2005, *4*, 27.

[30] De Marco, A., Schrödel, A., Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem.* 2005, *6*, 10.

[31] Ventura, S., Villaverde, A., Protein quality in bacterial inclusion bodies. *Trends Biotech.* 2006, *24*, 179–185.

[32] Gonzàlez-Montalbàn, N., Garcia-Fruitos, E., Ventura, S., Aris, A., Villaverde, A., The chaperone DnaK controls the fractioning of functional protein between soluble and insoluble cell fractions in inclusion body-forming cells. *Microb. Cell Fact.* 2006, *5*, 26.

[33] Gonzàlez-Montalbàn, N., Garcia-Fruitos, E., Villaverde, A., Recombinant protein solubility – Does more mean better? *Nat. Biotechnol.* 2007, *25*, 718–720.

[34] Garcia-Fruitos, E., Martinez-Alonso, M., Gonzàlez-Montalbàn, N., Valli, M. *et al*., Divergent genetic control of protein solubility and conformational quality in *Escherichia coli*. *J. Mol. Biol.* 2007, *374*, 195–205.

[35] Jevsevar, S., Gaberc-Porekar, V., Fonda, I., Podobnik, B. *et al*., Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. *Biotechnol. Prog*. 2005, *21*, 632–639.

[36] Singh, S. M., Panda, A. K., Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.* 2005, *99*, 303–310.

[37] Kirk, O., Borchert, T. V., Fuglsang, C. C. Industrial enzyme applications. *Curr. Opin. Biotechnol*. 2002, *13*, 345–351.

[38] van Beilen, J. B., Li, Z., Enzyme technology: An overview. *Curr. Opin. Biotechnol*. 2002, *13*, 338–344.

[39] Nilsson, M. R., Techniques to study amyloid fibril formation *in vitro*. *Methods* 2004, *34*, 151–160.

[40] Bondos, S. E., Methods for measuring protein aggregation. *Curr. Anal. Chem*. 2006, *2*, 157–170.
[41] Arrondo, J. L. R., Goni, F. M., Structure and dynamics of membrane proteins as studied by infrared spectroscopy. *Prog. Biophys. Mol. Biol.* 1999, *72*, 367–405.
[42] Dong, A., Huang, P., Caughey, W. S., Protein secondary structures in water from 2nd-derivative amide-I infrared spectra. *Biochemistry* 1990, *29*, 3303–3308.
[43] Natalello, A., Ami, D., Brocca, S., Lotti, M., Doglia, S. M., Secondary structure, conformational stability and glycosylation of a recombinant *Candida rugosa* lipase studied by Fourier-transform infrared spectroscopy. *Biochem. J.* 2005, *385*, 511–517.
[44] Ami, D., Bonecchi, L., Calì, S., Orsini, G., Tonon, G., Doglia, S. M., FT-IR study of heterologous protein expression in recombinant *Escherichia coli* strains. *Biochim. Biophys. Acta* 2003, *1624*, 6–10.
[45] Ami, D., Natalello, A., Gatti Lafranconi, P., Lotti, M., Doglia, S. M., Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. *FEBS Lett*. 2005, 579, 3433–3436.
[46] Barth, A., Zscherp, C., What vibrations tell us about proteins. *Q. Rev. Biophys.* 2002, *35*, 369–430.
[47] Natalello, A., Ami, D., Doglia, S. M., Protein aggregation studied in intact cells by Fourier transform infrared spectroscopy. In: Uversky, V. N., Permyakov, A. E. (Eds.), *Methods in Protein Structure and Stability Analysis: Vibrational Spectroscopy*, Nova Science, Hauppauge 2007, pp. 249–265.
[48] Chiti, F,. Dobson, C. M., Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 2006, *75*, 333–366.
[49] Zandomeneghi, G., Krebs, M. R. H., McCammon, M. G., Fandrich, M., FTIR reveals structural differences between native beta-sheet proteins and amyloid fibrils. *Protein Sci.* 2004, *13*, 3314–3321.
[50] Seshadri, S., Khurana, R., Fink, A. L., Fourier transform infrared spectroscopy in analysis of protein deposits. *Method Enzymol.* 1999, *309*, 559–576.
[51] Ami, D., Natalello, A., Zullini, A., Doglia, S. M., Fourier transform infrared microspectroscopy as a new tool for nematode studies. *FEBS Lett.* 2004, *576*, 297–300.
[52] Ami, D., Neri, T., Natalello, A., Mereghetti, P. *et al*., Embryonic stem cell differentiation studied by FT-IR spectroscopy. *Biochim. Biophys. Acta* 2007 (in press), DOI:10.1016/j.bbamcr.2007.08.003.
[53] Naumann, D., Helm, D., Labischinski,H., Microbiological characterizations by FT-IR spectroscopy. *Nature* 1991, *351*, 81–82.
[54] Fabian, H., Choo, L. P., Szendrei, G. I., Jackson, M. *et al*., Infrared spectroscopic characterization of Alzheimer plaques. *Appl. Spectrosc.* 1993, *47*, 1513–1518.
[55] Choo, L. P., Wetzel, D. L., Halliday, W. C., Jackson, M. *et al*., *In situ* characterization of beta-amyloid in Alzheimer's diseased tissue by synchrotron Fourier transform infrared microspectroscopy. *Biophys. J.* 1996, *71*, 1672–1679.
[56] Kretlow, A., Wang, Q., Kneipp, J., Lasch, P. *et al*., FTIR-microspectroscopy of prion-infected nervous tissue. *Biochim. Biophys. Acta* 2006, *1758*, 948–959.
[57] Gross-Selbeck, S., Margreiter, G., Obinger, C., Bayer, K., Fast quantification of recombinant protein inclusion bodies within intact cells by FT-IR spectroscopy. *Biotechnol. Prog.* 2007, *23*, 762–766.
[58] Orsini, F., Ami, D., Villa, A. M., Sala, G. *et al*., FT-IR microspectroscopy for microbiological studies. *J. Microbiol. Methods* 2000, *42*, 17–27.
[59] Umetsu, M., Tsumoto, K., Nitta, S., Adschiri, T. *et al*., Nondenaturing solubilization of beta2 microglobulin from inclusion bodies by L-arginine. *Biochem. Biophys. Res. Commun.* 2005, *328*, 189–197.
[60] Carrió, M., Gonzáles-Montalbán, N., Vera, A., Villaverde, A., Ventura, S., Amyloid-like properties of bacterial inclusion bodies. *J. Mol. Biol*. 2005, *347*, 1025–1037.

