Build-up curve processing.

- THE BUILD-UP CURVES PROCESSING IN GIFA
- THEORY
- PRACTISING THE BUILD-UP CURVES PROCESSING
- REFERENCES :

- Contents

THE BUILD-UP CURVES PROCESSING IN GIFA

In this chapter, the build-up curve processing command set available in Gifa is exposed. This set permits to calculate the distance from the results of a multi-exponential analysis (RELAX, DIST), and to evaluate the quality of this analysis (RCRYST). The processing of build-up curves is closely related to the Linear Prediction Package available in Gifa, and the manual related to the latter should be read before going into this one.

THEORY

- The multi-exponential expression of the NOE intensities
- Processing the complexe conjugate roots
- Checking the quality of the analysis
- Calculating the distances from the relaxation rates

The multi-exponential expression of the NOE intensities

One can express the nOe intensities obtained from a NOESY experiment, as a series of symmetric matrices $I(\tau_m)$, where $I_{ij}(\tau_m)$ holds the intensity of the cross peak between atoms i and j, for a NOESY mixing time of τ_m. The relaxation parameters for each pair of spins can be presented as the relaxation matrix Σ. Σ is bound to the nOe intensity matrix $I(\tau_m)$ by the relation :

$$I(\tau_m) = \exp(-\Sigma\tau_m) I_0 \quad (1)$$

where I_o is a diagonal matrix with elements equal to the equilibrium magnetizations of each respective spin.

If the complete experimental determination of the $I(\tau_m)$ matrix can be performed at a given τ_m, then the determination of Σ is straightforward. However this is seldom the case, due to overlaps in the 2D NOESY spectra.

The problem is then one of incompleteness of $I(\tau_m)$ and several procedures have been proposed to interpret quantitative NOESY information in terms of the relaxation matrix Σ. We will show here, that this incompleteness can be avoided if one takes a different approach to the problem.

The relaxation matrix is symmetric, provided we use the construction proposed by Olejniczak (Olejniczak, E. T. (1989) *J. Magn. Res.* **81**, 392-394). It can be diagonalized :

$$\Sigma = L\Lambda L^t \quad (2)$$

where Λ is a diagonal matrix, and I is thus computed by :

$$I = L\exp(-\Lambda\tau_m) L^t \quad (3)$$

From Eq. (2) we can rewrite each element of matrix Σ in the following way :

$$\Sigma_{ij} = \sum_k L_{ik} L_{jk} \lambda_k \quad (4)$$

so that each element of matrix I becomes :

$$I_{ij} = \sum_k L_{ik} L_{jk} \exp(-\lambda_k \tau_m) \quad (5)$$

where the L_{ik} is the (i,k) element of the matrix L, and λ_k is the K^{th} element of the diagonal matrix Λ. We can rewrite Eqs. (4) and (5) as follows :

$$I_{ij}(\tau_m) = \sum_k A_{ij}^k \exp(-\lambda_k \tau_m) \quad (6)$$

and : $$\Sigma_{ij} = \sum_k A_{ij}^k \lambda_k \quad (7)$$

where : $A_{k:ij} = L_{ik} L_{jk}$ (8)

The form of Eq. (6) proves that NOESY build-up curves are sums of exponentials. Eqs. (6) and (7) show that we can deduce the value of the relaxation matrix elements from the analytical form of the build-up curve.

The process we propose for computing the relaxation matrix elements from the NOESY experimental intensities is then, given an experimental nOe build-up curve $I_{ij}(\tau_m)$ measured between atoms i and j, to extract the parameters A_{ij}^k and λ_{ij}^k of the multi-exponential decay.

From the parameters A_{ij}^k and λ_{ij}^k it is then easy to obtain the values of the relaxation parameter Σ_{ij}. Subsequently, it is possible to extract distance information, by assuming a dynamic model for the molecule (for instance rigid spherical tumbling).

In Gifa, the LP-SVD method is used for multiexponential analysis of the curves, because it permits the separation of the signal parameters from additional parameters arising from the noise.

Processing the complexe conjugate roots

When the noise level is high, however, instability in the polynomial rooting may generate complex roots outside the unit-circle. In this case, these roots are found as complex conjugated doublets called "Froissard doublets". Such doublets lead to very inaccurate reconstruction of the respective relaxation matrix elements. To circumvent this problem, we propose the rotation of such doublets around the median point, before the amplitude estimate, in order to bring the roots back to the real axis.

A Froissard doublet is characterized by two related roots, located within the unit-circle, with opposite frequencies, opposite imaginary phases and the same amplitudes and damping factors:

$$Z_1 = e^{a+ib} \quad Z_2 = e^{a-ib} \quad (A1)$$

time dependence generated by such a pattern is:

$$ie^{(a+ib)t} + ie^{(a-ib)t} = ie^{at}(e^{-ibt} - e^{ibt}) \quad (A2)$$

$$= 2e^{at}\sin(bt) \sim 2bte^{at} \text{ if } bt << 1$$

the rotation leads to:

$$Z'_1 = e^{a+b} \quad Z'_2 = e^{a-b} \quad (A3)$$

which gives:

$$e^{(a+b)t} - e^{(a-b)t} = e^{(a+b)t}(1 - e^{-2bt}) \quad (A4)$$

$\sim 2bte^{at}$ if $bt << 1$ and $b << a$

The developments of the two expressions are very similar and equivalent for small values of bt. The operation permitting the replacement of Eq. (A2) by (A4) is equivalent to pivoting the root doublet by $\frac{\pi}{2}$ about the center of the doublet, bringing the roots back onto the real axis.

This operation does not modify the first order of the build-up curve reconstruction and can be safely used when the imaginary part of the complex root is small compared to the real part. This condition can be translated into geometric terms, by enforcing that the complex roots must lie within a cone of angular extent α.

In some cases however, particularly when the signal to noise ratio is very low, the LP-SVD analysis produces complex root doublets corresponding to higher frequencies for which the equations A2) and (A4) are no longer equivalent. In this case the fast frequencies detected in the signal are probably associated to the noise. We thus chose to ignore the related roots altogether for the amplitude reconstruction step. This decision was supported by the fact that the corresponding amplitudes, when computed, often appear to be at least one order of magnitude smaller than the other amplitudes.

Checking the quality of the analysis

From the multi-exponential analysis performed on the build-up curve as described, the relaxation parameter Σ_{ij} is computed as shown in Eq. (7). It is obvious that the accuracy of the determination of the relaxation parameter is critically related to the accuracy of the multi-exponential analysis. In order to check the quality of this analysis, for each build-up curve processed, an "R factor" can be computed,.much in the way X-ray crystallographers do.

$$R = \frac{\sum_k \left| x_k^{calc} - x_k^{exp} \right|}{\sum_k \left| x_k^{exp} \right|} \quad (10)$$

Here x_k^{calc} is the value of the calculated build-up curve reconstructed from the multi-exponential parameters, and x_k^{exp} is the "experimental" data, for k running on the different values of τ_m. We found the R factor to be more discriminant when computed only for small values of τ_m.

Calculating the distances from the relaxation rates

From the rates, the distances can be calculated, by assuming a dynamic model for the molecule. In Gifa, the rigid spherical tumbling model is available and permits to calculate the distances from the relaxation rates, in the following way. Having chosen a calibration rate Σ_0, corresponding to the known distance d_o, all other distances d_{ij} are obtained from the rates Σ_{ij} :

$$\frac{\Sigma_0}{\Sigma_{ij}} = \left(\frac{d_{ij}}{d_0}\right)^6 \quad (11)$$

PRACTISING THE BUILD-UP CURVES PROCESSING

- The Different Commands
- An example of build-up curve processing
  - Setting the reference distance and rate
  - Processing a build-up curve

A set of command as been implemented into Gifa in order to permit the processing as described above. The set-up is such that these commands should be used in conjonction with the Lineara Prediction Module of Gifa. The processing should thus be performed on a build-up curve, held into the regular 1D area of Gifa. This build-up curve should be such as obtained with a set of NOESY spectra, obtained for various value of the mixing time [tau]m, regularly sampled from 0 to Tmax. It should be noted that the value for $\tau_m = 0$ is assumed to be present. However it is rarely usefull to measure this value, and it is preferable to insert a 0 value as the $\tau_m = 0$ point of the build-up curve.

The Different Commands

Build-up curve analysis thus starts as a regular 1D LP-SVD analysis. Commands specific to the processing described above are the following :
RELAXRATE Permits to get the relaxation rate from the amplitude and the damping factors, which are obtained by the multi-exponential analysis of the build-up curve.
METH p Determine the hydrogen pair type. p is either 1, 2 or 3 for hydrogen-hydrogen, hydrogen-methyl or methyl-methyl pairs.
DIST Performs the distance calculation from the relaxation rate, using a reference distance, and the type of hydrogen pair observed.
CALIBDI dist_ref rate_ref Permits to define a reference rate, which corresponds to a reference distance, a the motion model considered.
RCRYST n Performs the computation of a "crystallographic factor", between the data obtained from the multi-exponential analysis and the initial data. This factor is calculated from the first n curve points.
SLOPE n Performs a least-square fit on the n first curve points, to determine his initial slope.
RTPIV [alpha] Performs the processing of the pairs of complexe conjugate roots, which can be obtained when solving the PE polynomial. The roots found inside the cone of extent [theta], are pivoted, those outside are removed: [alpha]= cotan([theta]). Works only with forward roots.
The following internal variables are available :

$RELAX $RCRYST $CALIBDI[] $DIST

An example of build-up curve processing

Setting the reference distance and rate

```
print"Calibration distance?" set dstcal = $_
print"Calibration relaxation rate?" set ratecal = $_

calibdi $dstcal $ratecal
```

Processing a build-up curve

```
print "Name of the build-up curve file?"
        set bldp = $_
print "Number of points on which the R-factor is computed?"
        set nb_pts_R = $_
print "Number of points on which initial-slope is computed?"
        set nb_pts_slop = $_
print "Order of the linear prediction analysis?"
        set ord = $_

read $bldp ; read data file
reverse chsize +1 reverse ; add the first (null) point
slope $nb_pts_slop dist ; determine the distance
; by initial slope method
chsize *2 swa ; makes the data complex by iterleaving zeros
; (the Gifa linear prediction package
; processes only complexe data).
```

order $ord ; defines the order of the linear ; prediction analysis
dt->svd % svd->ar 2 ; calculates the AR coefficients
rtclean 2 rtinv 2 rtfreq % 1 ; PE root processing
rt->pk % 1 ; calculate the damping factors and the ; amplitudes
relaxrate ; calculate the relaxation rate
dist ; calculate the distance

```
rcryst $nb_pts_R        ; calculate the crystallographic factor
```

REFERENCES :

- T.-E. Malliavin, M.-A. Delsuc and J.-Y. Lallemand, "Computation of Redfield matrix elements for incomplete NOESY data-sets",*J. Biomolec. NMR* **2**, 349-360 (1992)
- Reisdorf, C., Malliavin, T.-E. et Delsuc, M.-A. "Accurate estimation of inter-atomic distances in large proteins by NMR", *Biochimie* **74**, 809-813 (1992)

Build-up curve processing. - Contents

- THE BUILD-UP CURVES PROCESSING IN GIFA
- THEORY
  - The multi-exponential expression of the NOE intensities
  - Processing the complexe conjugate roots
  - Checking the quality of the analysis
  - Calculating the distances from the relaxation rates
- PRACTISING THE BUILD-UP CURVES PROCESSING
  - The Different Commands
  - An example of build-up curve processing
    - Setting the reference distance and rate
    - Processing a build-up curve
- REFERENCES :