

## Analysis of inelastic neutron scattering results on model compounds of the nitrogenous bases of the nucleotides

J TOMKINSON

The ISIS Facility, The Rutherford Appleton Laboratory, The Science and Technology Facilities Council, Chilton, Didcot, OX11 0QX, UK

E-mail. j.tomkinson@rl.ac.uk

**Abstract.** The role that model compounds can play in understanding the vibrational eigenvectors of molecules is discussed. Assigning the spectra of model compounds is of particular importance and the individual-scaling approach, that has been used with isolated molecule *ab-initio* calculations, is outlined. Special emphasis is given to recent work on assigning the spectra of three 5-6 heterobicyclic systems; indole, benzimidazole and isatin.

**Keywords.** Vibrational spectroscopy; nitrogenous bases; inelastic neutron scattering.

**PACS No.** 63.20

### 1. Introduction

Molecular vibrational spectroscopy offers access to structural details of complex biomolecular systems under realistic biological conditions. It achieves this using modern optical techniques, like resonance Raman scattering, to determine the subtle changes that occur in a few of their eigenvalues (the observed transition frequencies). However, the vast majority of the changes are ignored since the vibrational spectroscopy of large, and hence complex, systems is little understood. A deeper understanding of the molecular vibrations of larger systems must be available before this mass of data can be fully exploited. Inelastic neutron scattering (INS) spectroscopy offers new insights into the vibrations of molecules, which can be used to better understand the form of the molecular vibrations themselves, the eigenvectors.

One approach used in INS spectroscopy is to search for some aspects of the complex spectrum that may be recognisable in the spectra of simpler, structurally related, systems; their ‘model compounds’ [1]. However, to be useful, model compounds must be of a size with the structural components of the systems under study and currently, most cyclic model compounds studied by INS are small and non-heterocyclic. It seems reasonable to suggest that the spectra of 5-6

heterobicyclic molecules could form a reasonable base of model compounds to understand the eigenvectors of one interesting molecular system; the nitrogenous heterocyclic bases of the nucleotides.

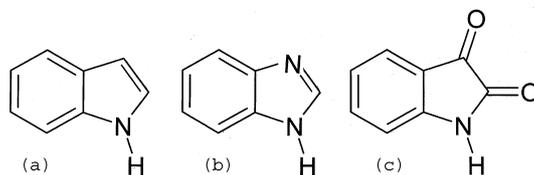
Low energy molecular vibrational eigenvectors involve atomic displacements over the molecule as a whole and the eigenvector is delocalised. This delocalisation induces modifications to the pattern of spectral intensities observed for, substantially, the same vibration when observed in different molecules. This applies to the results from any vibrational technique but is especially problematic for optical intensities, which are only indirectly related to the eigenvectors through the electro-optic parameters of the bulk material. Optical spectroscopy can avoid this problem by focussing on high-energy vibrations. Since optical techniques transfer insignificant momentum they have very favourable Debye–Waller factors at high energies and obtain good quality, high resolution results in this region. Here the vibrational eigenvectors are localised, involve only small groups of atoms within the molecule and so relate to only modest molecular distances. (Although subtle changes to these ‘group frequencies’ can be readily exploited.)

In contrast, INS intensities relate straightforwardly to atomic displacements and, so, suffer minimally from the delocalisation differences that affect an eigenvector common to several molecules. Thus, with INS there is a greater chance of exploiting the results of larger molecular model compounds than with optical techniques. Fortuitously, high resolution INS spectroscopy yields its most exploitable information at energy transfers below  $1500\text{ cm}^{-1}$ , because of unfavourable Debye–Waller factors at higher energies.

Here we shall discuss some recent work on assigning the eigenvectors of three 5-6 heterobicyclic molecules using the techniques of: single crystal orientation (indole) [2], deuteration (the case of benzimidazole) [2] and eigenvector elimination (isatin) [3] (see figure 1).

## 2. The INS technique

Apart from its ability to easily exploit the output from standard quantum mechanical calculations [1], INS spectroscopy can take advantage of two other practical techniques: single crystal orientation and deuteration. If all of the molecules in a sample can be oriented to place certain eigenvectors perpendicular to the momentum transfer vector,  $\mathbf{Q}$ , the intensities of these modes will disappear from the INS spectrum (or appear more strongly if oriented parallel to  $\mathbf{Q}$ ) [1]. Moreover, the strength of the INS signal is governed by the neutron scattering cross-section



**Figure 1.** The molecular structures of (a) indole, (b) benzimidazole and (c) isatin.

and eigenvectors involving hydrogen are always very clear in INS spectra. However, the total neutron scattering cross-section of deuterium (7.64 b) is an order of magnitude less than hydrogen (82 b) and any vibrations involving exclusively the vibration of a given hydrogen atom will disappear from the INS spectrum of its deuterated counterpart. The technique of eigenvector elimination is an approach to the spectral analysis that can be used, initially, to simplify the comparison of *ab-initio* calculation and INS observation. None of these techniques applies exclusively to INS spectroscopy and they can all be exploited, more or less, by optical spectroscopy but, when they are applicable, they are more easily used and produce more convincing results in INS.

### 3. Model compound exploitation

Even if we generate a basis of model compounds there is still the matter of how it can be exploited to understand new spectra. We need to compare the eigenvectors of the model compounds with those of the new system, extracting the signatures of promising eigenvectors from the spectra of the model compounds for comparison with the new spectra. We can attempt to compare the eigenvectors of different molecules by inspection but such comparisons are non-quantitative. Fortunately, the comparison can be made non-subjective by using the ViPA program [4]. Here, each eigenvector of the new molecule is expressed as a fraction of the eigenvectors in all vibrations of a model compound. As an example of the use of this program, in table 1 [2], the eigenvectors of indole are described in terms of those of its component molecules, benzene and pyrrole. The source of the eigenvector information is the *ab-initio* calculations that are currently at the heart of modern assignment schemes of all vibrational spectroscopy [1]. This is particularly appropriate in the case of INS, where calculated spectral results can be directly compared with the observed spectrum using ACLIMAX [5]. The *ab-initio* calculations discussed here were performed in GAUSSIAN98 [6] for isolated systems, the molecular geometries were constrained planar and optimised at the B3LYP, 6-31G\*\* level. The atomic displacement vectors of the harmonic vibrational frequencies were used as input to the ACLIMAX program.

Unlike optical spectroscopy, the INS technique expresses not only fundamentals but also all combination bands and overtones, and fortunately these are readily computed [1]. Critical to the appearance of these INS features is the reduced mass of the vibration involved. Generally, hydrogenous vibrations of low reduced mass produce strong fundamentals, combinations and overtone sequences in INS and also show significant H/D isotope shifts in optical spectroscopy [1].

### 4. Assignment by inspection, individual scaling

There are several approaches to handle the results of *ab-initio* calculations. Some workers prefer to leave them unchanged whilst others scale their *ab-initio* frequencies according to, more or less, complex schemes. When exploiting the results of isolated molecule calculations we prefer to use individual scaling to facilitate a visual comparison between the observed data and the final assignment scheme. Unfortunately, this is not a process that lends itself easily to the results from *ab-initio*

**Table 1.** Table of the eigenvectors of indole [2] given in terms of the modes of the parent cycles, benzene and pyrrole [4]. Where, Bz[ $x(y)$ ] implies that this indole mode has  $x\%$  of the benzene mode number  $y$  (after [10]); similarly for pyrrole [11], Py. Thus, the indole mode  $\nu_{31}$  (at  $937\text{ cm}^{-1}$ ) has 88% of the atomic displacements of the benzene mode  $\nu_{19}$  and no significant correlations to pyrrole.

A''			A'			Description
$\nu$	Obs.	G98	$\nu$	Obs.	G98	
42	239	222				Ring-ring anti-wag (Butterfly)
41	265	256				Ring-ring anti-twist
			29	397	410	Ring-ring anti-rock
40	429	445				Bz[91(20)]:Py[24(23), 14(19), 11(22)]
39	516	543				$\gamma$ -NH
			28	542	562	Bz[52(18)]:Py[16(7), 11(18)]
38	578	597				Bz[43(20), 26(8)]:Py[26(23), 12(12)]
37	608	625				Bz[28(4), 15(11)]:Py[40(23), 29(12), 14(11)]
			27	619	630	Bz[55(18)]:Py[23(20), 10(9)]
36	730	748				Bz[-]:Py[50(22), 19(11), 14(12)]
35	758	777				Bz[36(4), 40(11)]:Py[21(11)]
			26	758	781	Bz[40(18), 26(6)]:Py[20(6), 14(19), 12(9)]
34	–	819				Bz[45(11), 40(8)]:Py[15(12), 14(21)]
33	854	888				Bz[36(11), 18(19)]:Py[55(21)]
			25	867	903	Bz[44(7), 10(18)]:Py[25(8), 17(9)]
32	–	911				Bz[24(19), 12(11)]:Py[64(10), 10(21)]
			24	896	925	Bz[18(14), 9(6)]:Py[40(9), 32(20)]
31	937	972				Bz[88(19)]:Py[-]
30	971	1020				Bz[67(2), 30(19)]:Py[6(10)]
			23	1006	1043	Bz[57(14), 31(6)]:Py[7(7)]

calculations on extended crystals and it is often difficult, visually, to appreciate the quality of the assignment schemes based on such calculations.

The process of extracting individual scaling parameters pre-supposes significant knowledge of the character of each molecular transition as well as the calculated eigenvectors. Although, as in the case of indole below, this information might be accessible from INS results, it is more commonly provided from optical results. If this level of detail is not available, or if the system is of such low symmetry that no differentiation between transitions can be obtained, then the approach described below is little better than guesswork and should be avoided. However, when this level of detail is available, unambiguous assignments are generally possible across the vibrational range, except where several features are unresolved.

We begin at the lowest frequencies, and this has the advantage that fixing a fundamental also identifies its overtones and prevents strong overtones being mistaken for fundamentals. Taking any *ab-initio* frequency, we search for nearby observed transitions of similar INS intensity, that also possess equivalent characteristics. Thus, for example, if the optical character of the transition identifies it as coming from an out-of-plane mode it can only be associated with a calculated out-of-plane eigenvector. The observed band is thus assigned to a specific *ab-initio* eigenvector

and eigenvalue. We then scale this eigenvalue to equal the observed value, thus each transition has an individual scaling parameter. When this has been achieved for all the calculated transitions the scaled *ab-initio* spectrum should visibly match the observed spectrum; if not then the assignments are wrong.

## 5. The spectrometer

TOSCA is an indirect geometry time-of-flight spectrometer at the ISIS Facility, The Rutherford Appleton Laboratory, Chilton, UK [7]. The spectrometer has a low final neutron energy, very good spectral resolution ( $\Delta Et/Et \approx 1.3\%$ ) and covers a broad range, 20–4000  $\text{cm}^{-1}$ . Providing good energy resolution, at all energy transfers, sacrifices favourable Debye–Waller factors especially in the high energy transfer regime. As a consequence of its low final energy, and hence low final momentum, the momentum transfer vector,  $\mathbf{Q}$ , is essentially parallel to the incident beam for all energy transfers, at least  $>100 \text{ cm}^{-1}$ , and perpendicular to the face of the sample.

## 6. Discussion

### 6.1 Indole

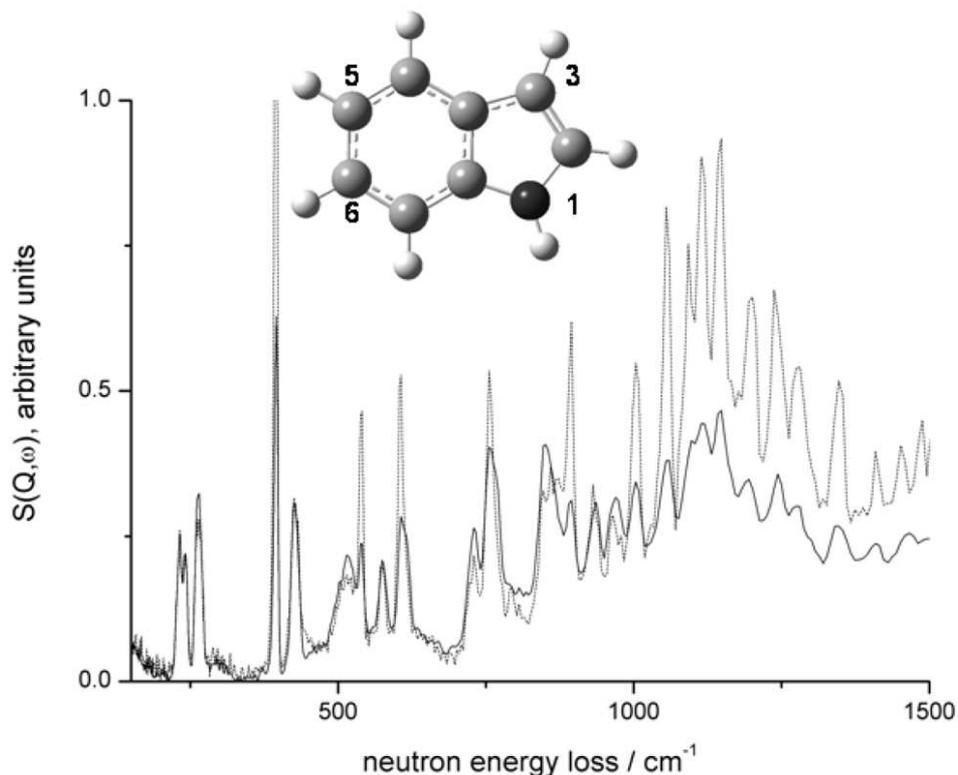
The observed INS spectrum of powdered and crystalline indole (20 K) is shown in figure 2 [2]. Although, in the crystalline spectrum, only eigenvectors in the molecular plane ( $A'$ ) should be INS-active a contaminating, pseudo-powder, spectrum sits below the single crystal spectrum. This reflects a poor alignment of the crystallites and, or multiple scattering. Both effects randomise the crystal alignment and produce the pseudo-powder spectrum [1]. However, where a band is stronger in the crystal spectrum it is from the oriented eigenvectors and the technique has clearly identified six of the seven  $A'$  modes below  $1050 \text{ cm}^{-1}$ .

A comparison of the INS of indole and  $d_2$ -indole did not yield such dramatic absences as seen below in the case of benzimidazole, except for the  $\gamma(\text{NH})$  mode, at  $516 \text{ cm}^{-1}$ , and another band about  $1109 \text{ cm}^{-1}$ . There are two reasons for this lack of contrast. First, deuteration of indole replaces only 2 of the 3 hydrogen atoms on the 5-ring and, second, very few of indole's eigenvectors involve, exclusively, the substituted atoms. The best that could be achieved was to observe significant intensity reductions in some modes. However, chasing intensity fluctuations can be a very subtle game in complex spectra and is usually not worth the effort.

Using the assignments from the INS spectra and the results of the *ab-initio* calculations the spectrum of indole was assigned by inspection, using ACLIMAX. The individually scaled quantum mechanical calculation was then compared with the observed INS spectrum in figure 3 [2].

### 6.2 Benzimidazole

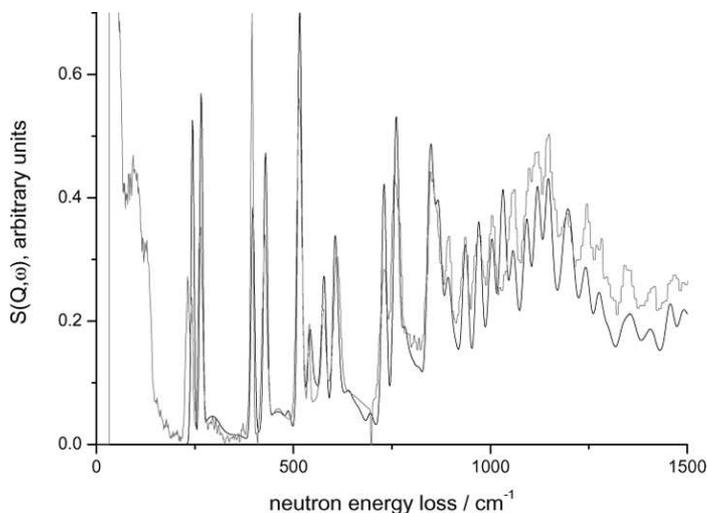
The INS spectra of benzimidazole and  $d_2$ -benzimidazole are shown in figure 4 [8]. Excellent agreement between the observed energies of optical and INS bands were obtained for both systems, except for some strong INS bands, with no optical counterparts, at  $918$  and  $1777 \text{ cm}^{-1}$ . The  $1777 \text{ cm}^{-1}$  transition appears in the



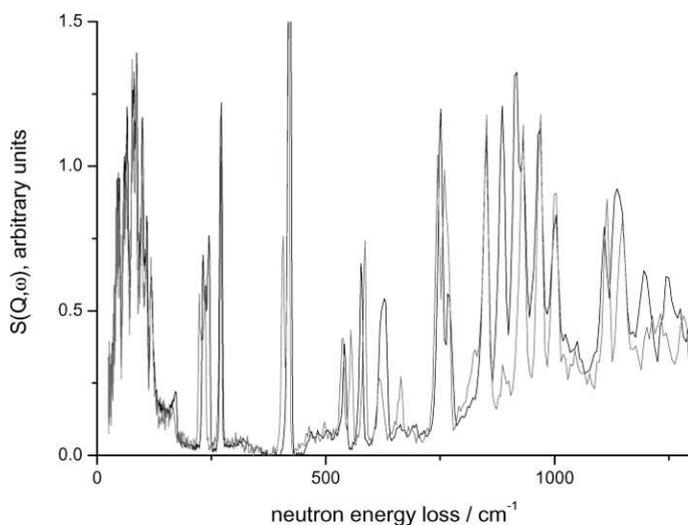
**Figure 2.** The INS spectra of powdered (solid line) and single crystalline (dashed line) indole, taken on TOSCA [2].

region of heavy atom stretches, like  $\nu(\text{C}=\text{O})$ , but it must arise from a hydrogen displacement to have any INS intensity, which immediately implies its assignment as an overtone [1]. Both the  $918\text{ cm}^{-1}$  fundamental, and its overtone, were absent from the INS spectrum of  $\text{d}_2$ -benzimidazole. The fundamental falls in the out-of-plane spectral mode region of the imidazole moiety of benzimidazole, which contains two eigenvectors,  $\gamma(\text{CH})_{\text{im}}$  and  $\gamma(\text{NH})$ . Furthermore, another benzimidazole band, at  $852\text{ cm}^{-1}$ , remained in the  $\text{d}_2$ -benzimidazole INS spectrum, at  $850\text{ cm}^{-1}$ , in the optical spectra it had disappeared.

The band at  $918\text{ cm}^{-1}$  was assigned to  $\gamma(\text{NH})$ . This out-of-plane mode was shown to be a light mass oscillator (ideal for generating overtones) in imidazole itself where it was observed in the same region. However, this high frequency assignment was in complete disagreement with published work. Optical studies placed  $\gamma(\text{NH})$  at  $628\text{ cm}^{-1}$  (observed in the INS, at  $632\text{ cm}^{-1}$ ) and compared it to  $\gamma(\text{NH})$  in indole. Both the INS and infra-red bands shifted on deuteration, to about  $555\text{ cm}^{-1}$ , but it retained significant INS intensity. This could only occur if its eigenvector involved displacements of the hydrogen atoms on the benzenoid moiety. It was shown to be an out-of-plane deformation of the heavy atoms of both rings, which the hydrogen atoms ‘rode’ [8]. The assignment of  $\gamma(\text{NH})$  at a relatively high frequency was



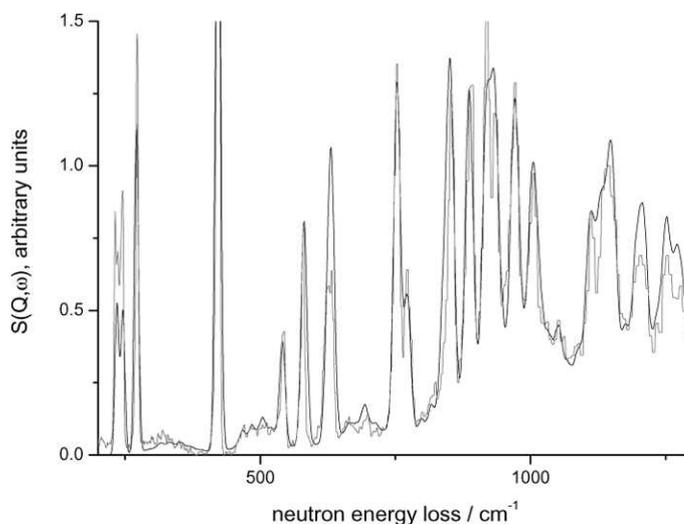
**Figure 3.** A comparison of the individually scaled *ab-initio* (black) and the observed spectra (gray) of indole, using ACLIMAX [2].



**Figure 4.** The INS spectra of benzimidazole (black) and  $d_2$ -benzimidazole (gray), 20 K, taken on TOSCA [8].

justified by *ab-initio* calculations on a dimer and is related to the strong NH...N bond found in benzimidazole [9]. In indole the strength of the hydrogen bonding is much weaker because the acceptor is the benzene ring of a neighbouring molecule, not an electronegative atom.

The assignment of  $\gamma(\text{CH})_{\text{im}}$  was less straightforward than  $\gamma(\text{NH})$  because it is mixed with a benzenoid mode,  $\gamma(\text{CH})_{\text{ben}}$ , producing in- and out-of-phase components, at 852 and 886  $\text{cm}^{-1}$ . Half of the intensity in these features came from



**Figure 5.** A comparison of the individually scaled *ab-initio* (black) and the observed spectra (gray) of benzimidazole, using ACLIMAX [8].

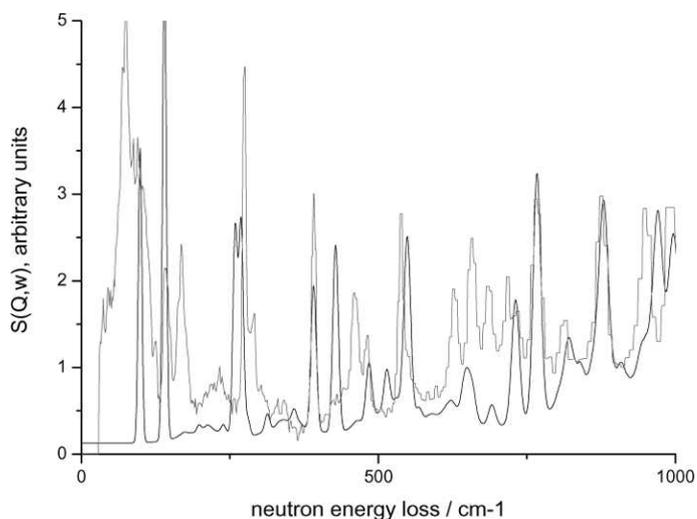
$\gamma(\text{CH})_{\text{im}}$  but in  $\text{d}_2$ -benzimidazole the  $\gamma(\text{CD})_{\text{im}}$  drops to a much lower frequency and the splitting collapses leaving a single feature, the intensity of which was exclusively from the benzenoid mode, and once more optical studies of the system had missed this subtlety.

Again using the assignments from the INS spectra and the results of the *ab-initio* calculations the spectrum of benzimidazole was assigned by inspection, using ACLIMAX. The individually scaled quantum mechanical calculation was then compared with the observed INS spectrum in figure 5 [8].

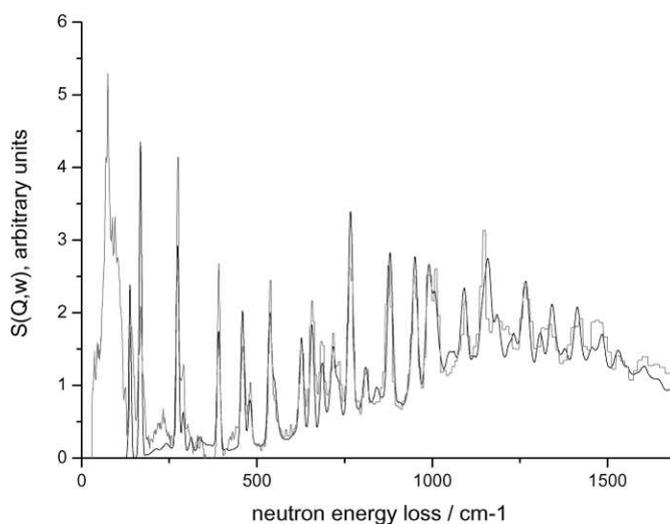
### 6.3 Isatin

In the case of isatin no deuteration or single crystal data was available and when the INS data were compared with the *ab-initio* results there was confusion in the mid-frequency range, about  $800 \text{ cm}^{-1}$  [3]. This was due to poorly calculated eigenvalues in this hydrogen bonded system. All N–H systems have a tendency to hydrogen bond in the solid state and this makes them difficult to treat as simple isolated molecules in any *ab-initio* calculation. Occasionally, they can be reasonably treated as dimers if the specifics of their intermolecular bonds are known. However, if they are multiply hydrogen bonded, or extended solids, like many oxides, the simple molecular concept fails and full crystal calculations are essential. That said, however, when the number of specific hydrogen bonds is low, N–H systems still, as seen above, retain relatively simple, or ‘pure’, eigenvectors. In the case of isatin, the molecules are isolated hydrogen bonded dimers and were treated as such in the *ab-initio* calculation.

To simplify the assignment process, recourse was to eliminate the eigenvectors of some modes. In this case it was readily achieved by defining the ACLIMAX



**Figure 6.** A comparison of the unscaled *ab-initio* (black) and the observed spectra (gray) of isatin, using ACLIMAX. Here, the cross-section of the N–H hydrogen atoms were set to zero (see text [3]).



**Figure 7.** A comparison of the individually scaled *ab-initio* (black) and the observed spectra (gray) of isatin, using ACLIMAX [8].

scattering cross-section of the N–H hydrogen atom, as zero. Removing the (N-) H cross-section eliminates any of its contributions from the *ab-initio* spectrum, in this case the ‘pure’  $\gamma(\text{NH})$  mode. This ‘eliminated eigenvector’ spectrum is shown in figure 6. The relationship of other calculated transitions to the observed spectrum is now much clearer.

Using the *ab-initio* calculations and ACLIMAX the INS spectrum was assigned, ignoring  $\gamma(\text{NH})$ . Eventually all but two bands, at 627 and 686  $\text{cm}^{-1}$ , had been assigned and what remained must be the eliminated  $\gamma(\text{NH})$  modes (two because the system is dimeric). The individually scaled quantum mechanical calculation was then compared with the observed INS spectrum in figure 7 [8].

## 7. Conclusion

The technique of INS spectroscopy offers significant hope that the mid-range vibrational spectra of complex molecules can be understood through the study of their model compounds. By constructing a knowledge base of those eigenvectors common to several members of this group of simple, geometrically related, systems, the vibrational assignments of the more complex molecules can be tackled with confidence. We have discussed some recent results obtained on 5-6 heterobicyclic systems that are the beginnings of a basis of model compounds for the study of the nitrogenous bases of the nucleotides and their related products.

## Acknowledgments

Here the authors would like to acknowledge the work of other authors involved in these studies; J Riesz and P Meredith [2]; M Bacci, M Piccollo and D Colognesi [3]; also my colleagues at ISIS, especially S F Parker and A J Ramirez-Cuesta. Thanks are due to the STFC, UK, for access to the ISIS Facility neutron beams.

## References

- [1] P C H Mitchell, S F Parker, A J Ramirez-Cuesta and J Tomkinson, *Vibrational spectroscopy with neutrons* (World Scientific Press, Singapore, 2005)
- [2] John Tomkinson, Jennifer Riesz, Paul Meredith and Stewart F Parker, *Chem. Phys.* **345**, 230 (2008)
- [3] J Tomkinson, M Bacci, M Piccollo and D Colognesi, *Vibrational spectroscopy*, *J. Phys. Chem. A*, accepted
- [4] A K Grafton and R A Wheeler, *Comput. Phys. Commun.* **113**, 78 (1998)
- [5] A J Ramirez-Cuesta, *Comput. Phys. Commun.* **157**, 226 (2004)
- [6] M J Frisch *et al*, *GAUSSIAN98*, Revision A 3, Gaussian, Inc., Pittsburgh (1998)
- [7] D Colognesi, M Celli, F Cilloco, R J Newport, S F Parker, V Rossi-Albertini, F Sacchetti, J Tomkinson and M Zoppi, *Appl. Phys.* **A74** (Suppl.), S64 (2002); also: <http://www.isis.rl.ac.uk/molecularSpectroscopy/tosca/index.htm>
- [8] J Tomkinson, *J. Phys. Chem.* **A112**, 6115 (2008)
- [9] C J Dik-Edixhoven, H Schent and H Van der Meer, *Cryst. Struct. Commun.* **2**, 23 (1973)
- [10] G Herzberg, *Molecular spectra and structure, II infrared and Raman spectra of polyatomic molecules* (Van Nostrand, New York, 1945) p. 118
- [11] D W Scott, *J. Mol. Spectrosc.* **37**, 77 (1971)  
see also T D Klots, R D Chirico and W V Steele, *Spectrochim. Acta* **A50**, 765 (1994) (table 4)