Article

A nomenclature and data model to describe NMR experiments

Rasmus H. Fogh^{a,†}, Wim F. Vranken^{b,†}, Wayne Boucher^a, Tim J. Stevens^a & Ernest D. Laue^{a,*}

^aDepartment of Biochemistry, University of Cambridge, Cambridge, CB2 1GA, UK; ^bMacromolecular Structure Database, European Bioinformatics Institute, Hinxton, Cambridge, CB10 1SD, UK

Received 17 April 2006; Accepted 9 August 2006

Key words: CCPN, data model, data standard, experiment nomenclature, experiment pathway, NMR

Abstract

Despite ongoing efforts in organising NMR information, there is no consistent and well-described generic standard for naming NMR experiments. The main reason for the absence of a universal naming system is that the information content of the coherence pathways is difficult to describe in full detail. We propose a system that describes the common and generic elements of the coherence pathways produced by pulse sequences. The system itself is formalised by an 'NMR experiment protocol' model, which is described in the Universal Modelling Language (UML) as part of the CCPN data model. Furthermore, normalized experiment names can be derived from this proposed model. We hope this article will stimulate discussion to organise the wealth of NMR experiments, and that by bringing this discussion into the public domain we can improve and expand our proposed system to include as much information and as many NMR experiments as possible.

Introduction

Since 1976, the application of Fourier transforms to NMR experiments has resulted in the development of a wealth of pulse sequences for multiple dimension NMR spectroscopy. Even for NMR experiments that have significantly different pulse sequences and different common names, the underlying coherence pathways are often very similar. However, despite some efforts in standardizing NMR experiment names in particular software packages such as SPINS (Baran et al., 2002), and in software provided by spectrometer manufacturers, there is no consistent and welldescribed standard for naming NMR experiments. Several IUPAC naming efforts in the NMR area (Markley et al., 1998; Harris et al., 2002) have also not addressed this issue. The main reason for the absence of a universal naming system is that the information content of the coherence pathways is difficult to describe in full detail. The relationships between the magnetisation transfers and different atom types involved in an NMR experiment can be very complex, and can only be well described in a system that allows a high degree of connectivity between the elements involved in a coherence pathway.

For example, some experiments, like the classical HNCOCA (Ikura et al., 1991; Grzesiek and Bax, 1992), transfer magnetization from the protein backbone hydrogen atom to the alphacarbon via the backbone carbonyl carbon atom, and then back again in reverse order. An unambiguous description of such a pathway has to clearly indicate that the initial atom involved in the pathway is the same one as the final one. Also, a

[†]Joint first authors

^{*}To whom correspondence should be addressed. E-mail: e.d.laue@bioc.cam.ac.uk

transfer pathway in an actual experiment often includes atoms that are not observed in the final spectrum, and these have to be distinguished from the detected atoms. Other problems involve splitting the pathway via different kinds of atoms (as in a combined ¹⁵N/¹³C HSQC NOESY (Boelens et al., 1994)), and identifying where special transfer types occur (like TOCSY (Braunschweiler and Ernst, 1983) or NOESY (Jeener et al., 1979)). The common name used for an experiment often does not encompass this information, e.g. the name 'HCCH-TOCSY' (Bax and Ikura, 1991; Olejniczak et al., 1992) is understood to be a 3D experiment with two proton dimensions and one carbon dimension, but the name does not in itself describe whether the carbon observed is the one before or after the TOCSY transfer.

Here, we propose a system that describes the common and generic elements of the coherence pathways produced by pulse sequences. This system does not describe all of the information from a pulse sequence in full detail, as much of the specifics are manufacturer dependent and too complex to standardize in a generic approach. The system itself is formalised by an 'NMR experiment protocol' model which is part of the CCPN data model (Fogh et al., 2002, 2005; Vranken et al., 2005), and normalized experiment names are derived from this proposed model. Experiments are organised into 'Experiment Prototypes' according to their magnetisation transfer path, which may well correspond to a 5D or 6D experiment. This is the level of abstraction you need to define and interpret peak assignments. Individual experiments are described and named separately, according to which nuclei are actually measured, reduced-dimensionality encoding, etc. Distinctions such as HSQC vs. HMQC transfer, constant vs. variable time, or coupled vs. decoupled acquisition are not taken into account. Individual experiments have a user-selected synonym which can be used to add more detail, e.g. with precise transfer types, or to make the name more intelligible, e.g. with a prefix to signify the overall dimensionality.

Materials and methods

The elements of the NMR reference experiment protocols are described as part of the CCPN data model (Fogh et al., 2002, 2005; Vranken et al.,

2005). Such a data model describes the organisation of information as classes of *objects*. Each class of object possesses a list of attributes and links to other objects, which may be of the same or a different class. The model is defined using UML (Unified Modelling Language), a computing industry standard for systems modelling which has the advantage of an associated graphical notation (see figures). The CCPN framework uses the generic UML model to derive program code for storing the relevant data (XML schemas, I/O libraries), and for manipulating the data in memory (Application Programming Interfaces (APIs) in Python, Java, and other languages). Data describing particular NMR experiments were entered manually into the CCPN framework using an *Experiment Prototype Editor*. This information was then stored in a well-defined and consistent way, using the data model, and can be readily accessed as reference data using the associated APIs.

Access to this reference data at the API level can be achieved on any platform that supports Python or Java. Moreover, the Experiment Prototype Editor is a Graphical User Interface (GUI) tool (Figure 1) that allows users to access the reference information and add their own reference experiment descriptions or change existing ones. This GUI depends on only Python and Tcl/Tk, and has already been run successfully on Irix, Linux, Mac OSX and Windows.

The CCPN Data Model itself is divided into 'packages', each of which describes one particular domain of information. The experiment descriptions from this article are stored in the NmrExpPrototype package. Because CCPN data model packages are connected to each other, this NmrExpPrototype package that contains the transfer pathway information is linked directly to the description in the Nmr package of a particular NMR experiment that resulted in a spectrum (Vranken et al., 2005) (see Figure 2). Accordingly, if you want to store a new NMR experiment in the Nmr package, you link this experiment to the correct reference experiment in the NmrExpPrototype package, and thus immediately get access to all the pathway information in that package (e.g. which types of atoms are connected by which type of transfer, ...). This also means that software that uses the data model directly like CcpNmr Analysis and the

Income of the local division of the local di	NmrExpPrototy											
#	Name	maxDim	Category	Synonym	Synonym Keywords Details isEditable ?				Deaue /			
20	H[N[CA]]	3	through-bond	HNCA	00	C C C D Later 22 Weatcoll =						
21	H[N[CO]]	3	through-bond	HNCO NMPEXPHOLOTYPE: 23: H[CA[CO]] V								
22	H[C[C]]	3	through-bond	HCC	-	Re	Experiments	-				
23	H[CA[CO]]	3	through-bond	HCACO	H		Name	numDims	1sRever	rsed T	isEditable	
24	H[C[N]]	3	through-bond	HCN	1		H[CA[CO]]	3	No		No	
25	H[N]_H.NOESY	3	through-space	15N HSQC-NOESY		Create New		Create Copy	Free	ze	Delete	
	Create New		Create Copy	F	-				1			
					-	Re	ExpDimRefs -					
	AtomSites	718	and and and and	10 may 70	Dim		expMeasuremen	t Splitting	Coupled	Constant	expSteps	
#	isotope	Name	Shift	Shift	1	1	1 Shift(H)	No		variable	e	
2	100		40,000	75.000		0	internet I	Delate Doff	in Plan	Periodo De	(EuroPhen Deef	
4	130	Ch ch	40.000	75.000		Create Nev		Desete nere	spom	Delete Pa	responser	
3	130		122,000	225,000								
					_			120120-02210-02210-001	E Page set a Fi	100 CO.		
		Log to t		0	d	_	Purge	RefExpDims	Close	Help		
Orea	te New - Type :	н⊽		Nmr	ExpProtot	ype	Purge 23: H[CA[CO]]	RefExpDims	Close	Help		
Oreal	te New - Type :	H V		Nmr	ExpProtot ExpGra	lype:	Purge 23: H[CA[CO]]	RefExpDims	Close teps	Help		
Orea	te New - Type : ExpMeasuremer	H V		Nmr	ExpProtot ExpGra	type: phs	Purge 23: H[CA[CO]] PeakSign	RefExpDims	Close teps StepNum	Help	Measurement	
Orea	te New - Type : ExpMeasuremen #	H V Its	ne e	atorsite	ExpProtot ExpGra # 1	lype: phs	Purge : 23: H[CA[CO]] PeakSign	RefExpDims	Close teps StepNum 1	Help	Measurement 1 Shift(H)	
Orea	te New - Type : ExpMeasuremen II	H V Its typ Shi	ne Ift	ator:Site	ExpProtot ExpGra # 1	type: phs	Purge 23: H[CA[CO]] PeakSign	RefExpDims	Close teps StepNum 1	Help	Measurement 1 Shift(H) 2 Shift(CA)	
Orea	te New - Type : ExpMeasuremen 1 2	H V Its typ Shi Shi	ne ft ft	atomSite 1(H) 2(CA)	ExpProtot ExpGra # 1	type: phs	Purge 23: H[CA[CO]] PeakSign	ExpS	Close teps StepNum 1 2	Help	Measurement 1 Shift(H) 2 Shift(CA)	
Orea	ExpMeasuremen # 1 2 3	H V Its typ Shi Shi	ne ft ft ft	atomSite 1(H) 2(CA) 3(CO)	ExpProtot ExpGra # 1	type: phs	Parge 23: H[CA[CO]] PeakSign Delete	ERefExpDins ExpS ExpS ExpS Create	Close	Help	Measurement 1 Shift(H) 2 Shift(CA) 7 Defet	
Orea	ExpMeasuremen # 1 2 3	H V Its Shi Shi	ne Ift Ift Ift	atomSite 1(H) 2(CA) 3(CO)	ExpProtot ExpGra # 1	bype: phs	Purge 23: H[CA[CO]] PeakSign Dolete	V ExpS H 2 Create	Close	Help	Measurement 1 Shift(H) 2 Shift(CA) r Defet	
Orea	te New - Type : ExpMeasuremen I 2 3	H V Its Shi Shi	ne ft ft ft	atorSite 1(H) 2(CA) 3(CO)	ExpProtot ExpGra # 1 Create Ne ExpT	type: phs sw	Purge 23: H[CA[CO]] PeakSign Delete	RefExpDims ExpS H 1 Create	Close	Help	Measurement 1 Shift(H) 2 Shift(CA) 7 Delet	
Orea	te New - Type ; ExpMeasuremen H 1 2 3	H V Its typ Shi Shi	ne ft ft ft	atorSite 1(H) 2(CA) 3(CO)	ExpProtot ExpGra # 1 Create He ExpT #	type: phs ew	Purg 23: H[CA[CO]] PeakSign Dulote Dulote	RefExpDims ExpS H 1 2 Create AtorSite1	Close teps StepNum 1 2 New Ato	Help	Measurement 1 Shift(H) 2 Shift(CA) T Delet transfe Toself	
Orea	te New - Type : ExpMeasuremen H 2 3	H V Its Shi Shi Shi	ne ft ft ft	atorisita 1(h) 2(CA) 3(CO)	ExpProtot ExpProtot ExpGra H 1 Create No ExpT H 1	lype: phs ew	Purg 23: H[CA[CO]] PeakSign Dolote Dolote Sters transfer Type onebond	RofExpDins ExpS H Create AtomSite1 1 (H)	Close teps StepNum 2 New Atc 2	Help ber Re-number mSite2 (CA)	Measurement 1 Shift(H) 2 Shift(CA) r Delet transfe ToSelf No	
Orea	te New - Type : ExpMeasuremen H 1 2 3	H V its Shi Shi	se ft ft	atomSite 1(H) 2(CA) 3(CO)	ExpProtot ExpProtot ExpGra H 1 Create No ExpT H 1	type: phs ew	Purge 23: H[CA[CO]] PeakSign Dolote fors transfer Type onebond Oreate New	AtomSite1	Close	Help	Measurement 1 Shift(H) 2 Shift(CA) r Delet transfi Toseif No ta	
Orea	te New - Type : ExpMeasurement H 1 2 3	H V Its Shi Shi Create New	pe ft ft ft ft starter	atorSite 1(H) 2(CA) 3(CO) xperiments	ExpProtot ExpOrate H Cruate Nr ExpT H 1	type: phs ew	PeakSign Delote Delote fors transfer Type onebond Create New	AtomSite1	Close	Help	Measurement 1 Shift(H) 2 Shift(CA) Defet transfe Toself No to	

Figure 1. The Experiment Prototype Editor.



Figure 2. Overview of the NmrExpPrototype package (light grey) and its connections to the Nmr package (dark grey). Both of these packages are part of the wider CCPN data model. In the UML notation used in this and other figures, boxes represent classes, while the connecting lines are links between them. Symbols besides the lines refer to the number of permitted objects: an ExpStep must be connected to exactly one ExpMeasurement, an ExpTransfer to exactly two AtomSites, an AtomSite may be connected to any number of ExpTransfers, and an ExpDim may be connected to a single RefExpDim. Lines with arrows represent one-way links, and lines with filled diamonds represent a 'part-of' relationship – e.g. an ExpMeasurement is and remains part of a specific NmrExpPrototype. Attributes (used in other figures) are written inside the boxes as 'name:' or 'name: = default-value'.

CcpNmr FormatConverter (Vranken et al., 2005) have immediate access to the reference experiment description data.

Results

Experiment naming convention

Because of the high level of organisation imposed by the NmrExpPrototype model, it is possible to derive standardised names for the pathway and the experiment in question. Full details of the classes of object and their connections are given in the 'Reference Experiment Data Model' section. In essence an NMR experiment is described by a series of magnetisation transfer steps between different atom sites, and some of these sites may have NMR measurements recorded.

For a 'straight-through' experiment with only one-bond or J-coupling transfer, we propose a name that is composed of the atom names of the relevant atom sites given in order, such as 'HNCAH', or 'HCACONH' (Ikura et al., 1990b; Clubb et al., 1992; Grezesiek and Bax, 1993a). For these examples 'H', 'N', 'CA', and 'CO' respectively represent hydrogen, nitrogen, alpha-carbon and carbonyl carbon atom sites.

Where the pathway involves a transfer step that is not a standard J-coupling we insert an underscore to indicate the other transfer type, and name it after a dot at the end of the experiment name. For example a two-dimensional hydrogen-hydrogen NOESY experiment would take the name 'H H.NOESY'. When combining J-coupling transfers with another kind of transfer, for example in the carbon TOCSY experiment commonly called HCCH-TOCSY (Bax et al. 1990a; Kay et al., 1990a; Ikura et al., 1991), we get the name 'HC CH.TOCSY', indicating that the TOCSY transfer is between the two carbon atom sites. If multiple other types of transfers occur, a dash separates them. For example 'H_H_H.NOESY-TOCSY' is a three-dimensional hydrogen experiment involving a NOESY transfer before a TOCSY transfer (Oschkinat, et al. 1988; Vuister et al., 1988).

'Out-and-back' magnetisation transfer is indicated by square brackets. For example, a protein backbone carbon experiment that goes from the amide hydrogen through the amide nitrogen and carbonyl carbon to an alpha-carbon and all the way back again is called 'H[N[CO[CA]]]'. 'Outand-back' steps can easily be combined with a non-J-coupling transfer: for example, we can describe the three-dimensional ¹⁵N HSQC involving a final ¹H NOESY step (Marion et al., 1989a, b; Zuiderweg and Fesik, 1989) as 'H[N]_H.NOESY'.

Measurements other than chemical shifts are given explicitly, as for example H[J(HH)], the 2D J-resolved experiment (Aue et al., 1976b). Where the magnetization may follow several different paths, they are given in curly braces. Mutually exclusive paths are separated by the logical OR sign '|', so that the pathway for a combined $^{15}N/^{13}C$ HSQC NOESY becomes H[{N|C}]_H. NOESY. Curly braces are also used to signify that two transfers happen simultaneously, in which case the alternatives are separated by a plus sign. An example would be the HNHA experiment used for measuring J(HHA) coupling constants (Vuister and Bax, 1993; Kuboniwa et al., 1994), which becomes H{[N]+[HA]}. When atom sites in the transfer path are never actually measured, their names are given in lower case. For example, 'H[C[co]]' is a two dimensional ¹³C HSQC (Müller, 1979; Bodenhausen and Ruben, 1980; Bax et al., 1990b; Norwood et al., 1990), selecting carbons bound to a carbonyl carbon. A further, more sophisticated example would be 'H[{n|c}]_H[{n|c}].NOESY', a two dimensional NOESY with 'labeled-to-labeled' selection for the adjacent carbon or nitrogen.

Selected/rejected experiments may also need a notation for the multiplicity of an atom site, which is appended in parenthesis, as for H[C[h(2n + 1)]] (2D ¹³C HSQC, selecting carbons bound to an odd number of protons), or H{[n(0)],[c(0)]}_H.TOCSY (2D TOCSY, selecting protons on unlabeled C/N). Finally, a few NmrExpPrototypes may require some text in addition to the fully systematic name in order to reflect important details. For instance the constant-time ¹³C HSQC where peak signs are determined by the number of ¹³C coupling partners is named 'ct.H[C[{c(2n)|c(2n + 1)}]]'.

The reference NMR experiments themselves are 'selected' from the full NMR experiment prototype description. Thus, experiments within the same prototype reflect the same pathway(s), but they differ according to which atom sites are being observed. The names of atom sites that are not detected in a particular reference NMR experiment (but are part of the pathway) then become lower case. For example, the protein backbone carbon experiment 'H[N[CO[CA]]]' prototype contains 'H[N[co[CA]]]' and 'H[n[CO[CA]]]' experiments. These are both part of the same pathway, but in the 'H[N[co[CA]]]' the carbonyl carbon shift is not measured and in the 'H[n[CO[CA]]]' the nitrogen shift is not measured. If the experiment is acquired in inverse order relative to the prototype, the name is reversed, so that e.g. a 'H[$\{N|C\}$] H.NOESY' becomes a 'H [{N|C}]H.NOESY'. In case of a classical HCCH carbon TOCSY, the difference between observing the carbon before or after the TOCSY transfer is given by the HC_cH.TOCSY and Hc_CH.TOCSY experiment names. RefExperiment names may contain extra text, to distinguish versions that are constant/variable time, coupled/decoupled, reduced-dimensionality (see below), etc.. More detailed and descriptive names can be entered in the RefExperiment synonym attribute.

By way of example a list of commonly used NMR experiment names and the corresponding pathway name as described above are given in Table 1.

Reference experiment data model

The model underlying the proposed experiment naming convention contains a fuller description of the pathways, atom sites, experiments and measurements. The information is divided into NmrExpPrototypes that each describe a magnetic transfer pathway. (See Figure 2. For a description of the UML conventions used in these figures, see the Materials and methods section.) Each NmrExpPrototype then corresponds to several RefExperiments. For example, the 'H[N[CO[CA]]]' NmrExpPrototype mentioned above has several associated RefExperiments which include not only the commonly used 3D 'H[N[CO[CA]]]', but also the 4D 'H[N[CO[CA]]]', the 3D 'H[n[CO[CA]]]' and the reduced-dimensionality

Table 1. Examples of commonly used experiment names and corresponding proposed systematic name

Common name	Systematic name
COSY (Aue et al., 1976a)	H H.COSY
DQF-COSY (Piantini et al., 1982; Rance et al., 1983)	H H.DQF–COSY
TOCSY	H_H.TOCSY
NOESY	H H.NOESY
DQ (Braunschweiler et al., 1983)	H[DQ(HH)]
J-resolved	H[J(HH)]
HSQC-T1	H[N][T1(H)]
HSQC-T1rho	H[N][T1rho(H)]
HSQC-T1zz	H[N][T1zz(HN)]
NOESY-TOCSY	H_H_H.NOESY-TOCSY
13C HSQC/HMQC	H[C]
13C HMBC (Summers and Bax, 1986)	H_[C].Jmultibond
Aromatic-selective 13C HSQC (Grzesiek and Bax, 1993a)	H[C[caro]]
CO-selective 13C HSQC (Grzesiek and Bax, 1993a)	H[C[co]]
HCCH-COSY	HCCH
HCCH–TOCSY	HC_CH.TOCSY
HCCONH (Kay et al., 1990b)	HCCONH
HCC-TOCSY-CONH (Campbell-Burk et al., 1992;	HC_cCONH.TOCSY
Grzesiek et al., 1993; Logan et al., 1993;	
Lyons et al., 1993; Montelione et al., 1992)	
13C HSQC-TOCSY (Marion et al., 1989b)	H[C]_H.TOCSY
HCACO (Grzesiek and Bax, 1993b)	H[CA[CO]]
HCbNb (Sklenar et al., 1993a)	H[Cbase[Nbase]]
HACACON (Kay et al., 1990b; Powers et al., 1991; Palmer et al., 1992)	H[CA[CO[N]]]
HACANH (Ikura et al., 1991)	HA[CA[N[H]]]
HNCACB (Wittekind and Mueller, 1993)	H[N[CA[CB]]]
HCsNbCbH (Farmer et al., 1993; Sklenar et al., 1993b)	HCsugarNCbaseH
HBCB/HACANH (Wang et al., 1994)	H{CA Cca}NH
15N/13C HSQC-TOCSY	H[{N C}]_H.TOCSY
13C HSQC-NOESY (Marion et al., 1989b)	H[C]_H.NOESY
13C,13C HSQC-NOESY-HSQC (Frenkiel et al., 1990; Ikura et al., 1990a)	H[C]_H[C].NOESY
13CH \rightarrow 12C HSQC-NOESY (Otting and Wuthrich, 1989)	H[C]_H[c(0)].NOESY
13CH HSQC \rightarrow 12CH/14NH NOESY (Otting and Wuthrich, 1989)	$H[C]_H{[c(0)] + [n(0)]}.NOESY$
13CH HSQC \rightarrow 13CH/15NH NOESY (Otting and Wuthrich, 1989)	$H[C]_H[{c n}].NOESY$
15N/13C HSQC-NOESY	$H[\{n c\}]_H.NOESY$
13 CH/15NH \rightarrow H NOESY (Otting and Wuthrich, 1989)	H[{n c}]_H.NOESY
13 CH/15NH \rightarrow 12CH/14NH NOESY (Otting and Wuthrich, 1989)	$H[\{n c\}]_H\{[n(0)]+[c(0)]\}.NOESY$

3D 'H[N[CO[CA]].CO splits CA'. This arrangement reduces the duplication of information.

The NmrExpPrototype class is shown in detail in Figure 3. The attributes for naming the pathway are name for its systematic name, and a common name (synonym) as an alternative to the systematic name. The *category* attribute serves to organise NmrExpPrototypes into 'through-bond', 'throughspace', 'measurement', or 'other', and the keywords and details attributes can be used to enter additional information. AtomSites describe the atoms that are involved in the magnetisation transfer, each with an isotopeCode (e.g. 'C'), and a name (e.g. 'CA'). The minShift and maxShift attributes serve to distinguish, for example, CO (155-225 ppm) from CA (40-75 ppm) from plain C (unrestricted). The remaining attributes are used to specify the numbers of atoms for experiments that select on the number of coupling partners. It is possible to specify, for example, 'exactly 1 proton', '2-3 protons', 'zero protons', or 'an odd number of protons'. The ExpTransfer class connects AtomSites to each other, and simply holds the transferType attribute and a link to the AtomSites in question. Note that an ExpTransfer tells you that there must be, for example, a J-coupling between the two AtomSites – whether it is used for magnetisation transfer, selection/rejection or something else is another matter. The



Figure 3. Detail of the NmrExpPrototype class and its connections to the ExpGraph, AtomSite and ExpTransfer classes.

possible transferTypes are 'NOESY' (which includes ROESY), 'TOCSY', 'onebond', 'Jcoupling', and 'Jmultibond' (which specifically excludes onebond coupling). The transferToSelf attribute records whether the magnetisation may remain on the original nucleus, giving rise to a diagonal peak. Finally, the ExpGraph class connects the ExpTransfers together. Most NmrExpPrototypes have only one ExpGraph, but some experiments involve separate parallel magnetisation transfer pathways, like the ${}^{13}C/{}^{15}N$ HSQC–NOESY, or the HNCA/CB experiment. In those cases the Exp-Graph selects the ExpTransfers and (implicitly) AtomSites that belong together. The peakSign attribute gives the relative sign of the resulting peak (see also Figure 4).

The actual quantities being measured are given in the ExpMeasurement class, shown in Figure 4. Most experiments have chemical shifts on the axes, but the *measurementType* may be any of 'Shift', 'ShiftAnisotropy', 'Jcoupling', 'Rdc', 'TROSY', 'Dipolar coupling', 'MQshift', 'T1', 'T2', 'T1rho', or 'T1zz'. Where the measurementType is 'Shift', 'TROSY' 'ShiftAnisotropy', 'T1', 'T2', or 'T1rho', the ExpMeasurement is linked to a single Atom-Site, showing which shift is being measured. Measurements of 'Jcoupling', 'Rdc', 'Dipolar coupling' or 'T1zz' are linked to two AtomSites. 'MQ shifts', finally, may be linked to as many AtomSites as needed to specify the multiple quantum magnetisation. To allow for situations like 3-spin single quantum magnetisation, the atom-SiteWeights attribute gives the weights for each contribution from each AtomSite. The ExpSteps



Figure 4. Detail of the ExpMeasurement class and its connections to the NmrExpPrototype, AtomSite, ExpStep and ExpGraph classes.

give the order in which the various ExpMeasurements are accessed, corresponding to the order of traversal of the AtomSite graph. An ExpMeasurement may appear several times – in e.g. a 2D HSQC the order of steps would be ¹H shift– ¹⁵N shift– ¹H shift.

The RefExperiment (see Figure 5) describes which of the ExpMeasurements available from the magnetisation transfer path are actually being recorded in a particular experiment. The *name* attribute is set to the systematic name as described in the Naming convention section, and the synonym allows for a more detailed or human-readable alternative name. The isReversed attribute signifies that the entire experiment is acquired backwards relative to the order in which the RefExpDim, ExpSteps etc. occur. This allows, for example, HNCAH and HCANH experiments to share a single NmrExp-Prototype. Each RefExpDim represents a dimension in the experiment. Finally, each RefExpDimRef identifies an ExpMeasurement that is being recorded in this particular dimension. In most cases there will be only one, but multiple RefExpDimRefs can appear for a RefExpDim for two reasons. It may be that several measurements are being recorded in parallel on the same axis, like for a ${}^{13}C/{}^{15}N$ HSQC–NOESY, or it may be that a measurement is being recorded as a splitting, like for a reduced-dimensionality experiment. The coupledIsotopeCodes attribute describes if couplings to protons, ¹³C etc. are active, and the *constantTime* attribute distinguishes between constant-time, variable-time, and mixed acquisition. The link to



Figure 5. The RefExperiment class and its connections to RefExpDim and RefExpDimRef.

ExpStep allows one to specify which of several possible steps is actually being used for recording the measurement – it is normally left unset.

As mentioned in the Materials and methods section, an Experiment that is created in the Nmr package can be directly linked to a RefExperiment reference experiment in the NmrExpPrototype package, and all the relevant information about experimental transfers, measurements and atom sites can then be accessed (see Figure 2). This should provide essential information for automatic procedures that perform, for example, chemical shift assignments: instead of depending on the name of an experiment and specifying which dimension contains what information, one can directly write generic routines that access the type of atoms and transfers involved from the reference experiment data. This means that information from experiments that is relevant for a specific software package can be automatically identified and used.

The full hyperlinked table of currently available NmrExpPrototypes and RefExperiments is available online at http://www.ebi.ac.uk/msd-srv/ docs/NMR/NMRExpPrototype/nmrExpPrototype. html. The list of common and systematic names for the current reference data is also available as supplementary material. Both the CcpNmr Analysis and FormatConverter programs have a selection menu where reference experiments can be selected from this list when creating an experiment in the Nmr package. The new Nmr experiment is set up using the reference experiment information, and links are automatically created between them so the reference experiment information is accessible.

Finally, the *Experiment Prototype Editor* can be used to extend the currently available list of NmrExpPrototypes and RefExperiments. Existing NmrExpPrototypes can be copied to serve as a template, or an entirely new NmrExpPrototype can be created. First, the relevant AtomSites and ExpMeasurements should be created, followed by the required ExpGraph, ExpTransfer and ExpSteps. The relevant RefExperiments with RefExpDims and RefExpDimRefs can then be selected from this information. It is usually easiest to start by creating the highest dimensionality RefExperiment, and make the others by copying and editing. The existing NmrExpPrototypes can be viewed within the editor or on the website and serve as examples when creating a new NmrExp-Prototype.

If any new NmrExpPrototypes are required or created, we will include them in the standard reference data included with the CCPN software.

Conclusion

We have described a system to organise, name and store reference data for most commonly used NMR pathways and experiments. Within the current system we have tried to describe as many common experiments as possible, and have included more exotic ones as proof of concept. Further details about the pathways and experiments could in principle be included, but we have chosen a solution that is a trade-off between trying to keep the experiment description relatively simple, while at the same time including as much information as possible.

The reference information for NMR experiments we described in this model is an important part of the CCPN effort. If NMR experiments stored in the CCPN framework are linked to the correct reference experiments, it becomes possible to write software routines that extract generic information about connections between spectral dimensions from the reference data. This can be extremely useful for automatic assignment routines, as the software is then no longer restricted to particular stored experiment names or connections, but can directly access the relevant information.

A final area where this information can be used is in ontologies that include NMR experiment descriptions. The standardised names should be useful to provide more detailed descriptors for the NMR experiments.

We hope this article will stimulate discussion to organise the wealth of NMR experiments. The model we propose is not necessarily the final answer, and we hope that by bringing this discussion into the public domain we can improve and expand it to include as much information and as many cases as possible.

Electronic supplementary material is available to authorized users in the online version of this article at http://www.dx.doi.org/10.1007/s10858-006-9076-z and is accessible for authorized users.

Acknowledgements

The authors thank Daniel Nietlispach for very useful discussions to steer the model in the right direction. We are grateful for support from the Biochemistry and Biological Sciences Research Council, UK (Grants 8/B13014 and 8/B20259), and the European Community "Quality of Life and Management of Living Resources" program (Grants QLRI-CT-2001-00015 "TEMBLOR" and QLG2-CT-2000-01313 "NMRQUAL").

References

- Aue, W.P., Bartholdi, E. and Ernst, R.R. (1976a) J. Chem. Phys., 64, 2229–2246.
- Aue, W.P., Karhan, J. and Ernst, R.R. (1976b) J. Chem. Phys., 64, 4226–4227.
- Baran, M.C., Moseley, H.N., Sahota, G. and Montelione, G.T. (2002) J. Biomol. NMR, 24, 113–121.
- Bax, A. and Ikura, M. (1991) J. Biomol. NMR, 1, 99-104.
- Bax, A., Clore, G.M., Driscoll, P.C., Gronenborn, A.M., Ikura, M. and Kay, L.E. (1990a) J. Magn. Reson., 87, 620–627.
- Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990b) J. Magn. Reson., 86, 304–318.
- Bodenhausen, G. and Ruben, D.J. (1980) Chem. Phys. Lett., 69, 185–189.
- Boelens, R., Burgering, M., Fogh, R.H. and Kaptein, R.E. (1994) *J. Biomol. NMR*, **4**, 201–213.
- Braunschweiler, L., Bodenhausen, G. and Ernst, R.R. (1983) *Mol. Phys.*, 48, 535–560.
- Braunschweiler, L. and Ernst, R.R. (1983) J. Magn. Reson., 53, 521–528.
- Campbell-Burk, S.L., Domaille, P.J., Starovasnik, M.A., Boucher, W. and Laue, E.D. (1992) J. Biomol. NMR, 2, 639–646.
- Clubb, R.T., Thanabal, V. and Wagner, G. (1992) J. Biomol. NMR, 2, 203–210.
- Farmer, B.T. II, Mueller, L., Nikonowicz, E.P. and Pardi, A. (1993) *J. Am. Chem. Soc.*, **115**, 11040–41.
- Fogh, R., Ionides, J., Ulrich, E., Boucher, W., Vranken, W., Linge, J.P., Habeck, M., Rieping, W., Bhat, T.N., Westbrook, J., Henrick, K., Gilliland, G., Berman, H., Thornton, J., Nilges, M., Markley, J. and Laue, E. (2002) *Nat. Struct. Biol.*, 9, 416–418.
- Fogh, R.H., Boucher, W., Vranken, W.F., Pajon, A., Stevens, T.J., Bhat, T.N., Westbrook, J., Ionides, J.M. and Laue, E.D. (2005) *Bioinformatics*, 21, 1678–1684.
- Frenkiel, T., Bauer, C., Carr, M.D., Birdsall, B. and Feeney, J. (1990) J. Magn. Reson., 90, 420–425.
- Grzesiek, S., Anglister, J. and Bax, A. (1993) *J. Magn. Reson. B*, **101**, 114–119.
- Grzesiek, S. and Bax, A. (1992) J. Magn. Reson., 96, 432-440.
- Grzesiek, S. and Bax, A. (1993a) J. Biomol. NMR, 3, 185-204.
- Grzesiek, S. and Bax, A. (1993b) J. Magn. Reson. B, 102, 103– 106.
- Harris, R.K., Becker, E.D., Cabral de Menezes, S.M., Goodfellow, R. and Granger, P. (2002) Concepts Magn. Reson., 14, 326–346.

- Ikura, M., Bax, A., Clore, G.M. and Gronenborn, A.M. (1990a) J. Am. Chem. Soc., 112, 9020–9022.
- Ikura, M., Kay, L.E. and Bax, A. (1990b) Biochemistry, 29, 4659–4667.
- Ikura, M., Kay, L.E. and Bax, A. (1991) J. Biomol. NMR, 1, 299-304.
- Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) J. Chem. Phys., 71, 4546–4553.
- Kay, L.E., Ikura, M. and Bax, A. (1990a) J. Am. Chem. Soc., 112, 888–889.
- Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990b) J. Magn. Reson., 89, 496–514.
- Kuboniwa, H., Grzesiek, S., Delaglio, F. and Bax, A. (1994) J. Biomol. NMR, 4, 871–878.
- Logan, T.M., Olejniczak, E.T., Xu, R.X. and Fesik, S.W. (1993) J. Biomol. NMR, 3, 225–231.
- Lyons, B.A., Tashiro, M., Cedergren, L., Nilsson, B. and Montelione, G.T. (1993) *Biochemistry*, **32**, 7839–7845.
- Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989a) J. Am. Chem. Soc., 111, 1515–1517.
- Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989b) *Biochemistry*, 28, 6150–6156.
- Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wuthrich, K. (1998) J. Mol. Biol., 280, 933–952.
- Montelione, G.T., Lyons, B.A., Emerson, S.D. and Tashiro, M. (1992) J. Am. Chem. Soc., 114, 10974–10975.
- Müller, L. (1979) J. Am. Chem. Soc., 101, 4481-4484.
- Norwood, T.J., Boyd, J., Heritage, J.E., Soffe, N. and Campbell, I.D. (1990) J. Magn. Reson., 87, 488–501.
- Olejniczak, E.T., Xu, R.X. and Fesik, S.W. (1992) J. Biomol. NMR, 2, 655–659.
- Oschkinat, H., Griesinger, C., Kraulis, P.J., Sorensen, O.W., Ernst, R.R., Gronenborn, A.M. and Clore, G.M. (1988) *Nature*, **332**, 374–376.

- Otting, G. and Wuthrich, K. (1989) J. Magn. Reson., 85, 586-594.
- Palmer, A.G. 3rd, Fairbrother, W.J., Cavanagh, J., Wright, P.E. and Rance, M. (1992) J. Biomol. NMR, 2, 103–108.
- Piantini, U., Sorensen, O.W. and Ernst, R.R. (1982) J. Am. Chem. Soc., 104, 6800–6801.
- Powers, R., Gronenborn, A.M., Marius Clore, G. and Bax, A. (1991) J. Magn. Reson., 94, 209–213.
- Rance, M., Sorensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. and Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, 117, 479–485.
- Sklenar, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1993a) J. Biomol. NMR, 3, 721–727.
- Sklenar, V., Peterson, R.D., Rejante, M.R., Wang, E. and Feigon, J. (1993b) J. Am. Chem. Soc., 115, 12181–82.
- Summers, M.F. and Bax, A. (1986) J. Am. Chem. Soc., 108, 2093–2094.
- Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J. and Laue, E.D. (2005) *Proteins*, **59**, 687–696.
- Vuister, G.W. and Bax, A. (1993) J. Am. Chem. Soc., 115, 7772–7777.
- Vuister, G.W., Boelens, R. and Kaptein, R. (1988) J. Magn. Reson., 80, 176–185.
- Wang, A.C., Lodi, P.J., Qin, J., Vuister, G.W., Gronenborn, A.M. and Clore, G.M. (1994) J. Magn. Reson. B, 105, 196– 198.
- Wittekind, M. and Mueller, L. (1993) J. Magn. Reson. B, 101, 201–205.
- Zuiderweg, E.R. and Fesik, S.W. (1989) *Biochemistry*, 28, 2387–2391.