Journal of Magnetic Resonance **154,** 28–45 (2002)
doi:10.1006/jmre.2001.2454, available online at http://www.idealibrary.com on IDEAL®

Specification and Visualization of Anisotropic Interaction Tensors in Polypeptides and Numerical Simulations in Biological Solid-State NMR

Mads Bak, Robert Schultz, Thomas Vosegaard, and Niels Chr. Nielsen[1]

Laboratory for Biomolecular NMR Spectroscopy, Department of Molecular and Structural Biology, University of Aarhus, DK-8000 Aarhus C, Denmark

Received May 12, 2001; revised October 3, 2001; published online November 29, 2001

Software facilitating numerical simulation of solid-state NMR experiments on polypeptides is presented. The Tcl-controlled SIMMOL program reads in atomic coordinates in the PDB format from which it generates typical or user-defined parameters for the chemical shift, *J* coupling, quadrupolar coupling, and dipolar coupling tensors. The output is a spin system file for numerical simulations, e.g., using SIMPSON (Bak, Rasmussen, and Nielsen, *J. Magn. Reson.* 147, 296 (2000)), as well as a 3D visualization of the molecular structure, or selected parts of this, with user-controlled representation of relevant tensors, bonds, atoms, peptide planes, and coordinate systems. The combination of SIMPSON and SIMMOL allows straightforward simulation of the response of advanced solid-state NMR experiments on typical nuclear spin interactions present in polypeptides. Thus, SIMMOL may be considered a "sample changer" to the SIMPSON "computer spectrometer" and proves to be very useful for the design and optimization of pulse sequences for application on uniformly or extensively isotope-labeled peptides where multiple-spin interactions need to be considered. These aspects are demonstrated by optimization and simulation of novel DCP and C7 based 2D N(CO)CA, N(CA)CB, and N(CA)CX MAS correlation experiments for multiple-spin clusters in ubiquitin and by simulation of PISA wheels from PISEMA spectra of uniaxially oriented bacteriorhodopsin and rhodopsin under conditions of finite RF pulses and multiple spin interactions.

INTRODUCTION

Numerical simulations play an increasingly important role in the development and application of solid-state NMR methods for determination of the structure and dynamics of biological macromolecules immobilized by size, aggregation, or membrane association. This is ascribed to the fact that most solid-state NMR experiments used for this purpose strongly depend on manipulation of anisotropic interactions to obtain evolution under well-defined isotropic or anisotropic parts of the Hamiltonian. Based on advanced RF irradiation schemes, often in concert with sample spinning, a large number of pulse sequences have been devised to accomplish specific coherence transfers or measurement of specific structural parameters. Typically, these elements have been developed on basis of one- or two-spin systems using analytical tools such as average Hamiltonian theory (*1, 2*) to tailor the effective Hamiltonian to the appropriate form. Subsequently, the elements have been tested by numerical simulations and by experiments on model systems and finally verified in real peptide applications either directly or as elements in more advanced pulse schemes. Thus, in the development process numerical simulations have primarily been used for verification while they so far have only been used sparsely for direct design of pulse sequences (*3, 4*). In contrast, numerical simulations in combination with iterative fitting are regarded as being almost indispensable for the extraction of structural parameters from experimental spectra (*5–15*).

Considering the increasing use of uniformly or extensively $^{13}C,^{15}N$-labeled samples, it becomes exceedingly important that the pulse sequence elements work appropriately in multiple-spin systems with characteristics potentially being far from the simple one- or two-spin cases typically considered in the design of these elements. Under multiple-spin conditions vital coherences may leak to undesired spins which may cause a significant reduction in the sensitivity, wrong assignment of multiple-dimensional spectra, and wrong interpretations of anisotropic interactions in terms of structure and dynamics. Obviously, this problem may scale dramatically with the number of dimensions and coherence transfer steps involved in the pulse sequence. This is an important issue since current remedies to the resolution problem of biological solid-state NMR, in addition to uniformly labeled samples, appear to involve increasingly sophisticated combinations of pulse sequence building blocks in multiple-dimensional experiments. An important ingredient in the solution of this problem may be to investigate in detail the performance of the available pulse schemes on multiple-spin systems closely reflecting the conditions in relevant peptide structures. This may provide optimized experimental procedures, motivate the design of new procedures on a multiple-spin basis, or result in the recommendation of alternative isotope labeling procedures (*16*) being optimally compatible with state-of-the-art solid-state NMR technology.

[1] To whom correspondence should be addressed. E-mail: ncn@imsb.au.dk.

1090-7807/02 $35.00

Facing the facts that (i) analytical pulse sequence evaluations become exceedingly difficult in multiple-spin cases, (ii) it is necessary to consider the combined action of the various elements forming the full pulse sequence, and (iii) experimental evaluations require large amounts of spectrometer time and a variety of samples with different structure/labeling, it appears reasonable to conduct such evaluations numerically. Preferably this should be accomplished using software which allows the easy interchange of pulse sequences *and* multiple-spin systems corresponding to different structures. The recently introduced SIMPSON "computer spectrometer" (*17*) serves most of the needs for efficient evaluation of advanced multiple-pulse sequences on multiple-spin systems. In this paper we present a program, SIMMOL, which enables straightforward establishment of the anisotropic interaction tensors required for the evaluation of pulse sequences on relevant peptide structures using SIMPSON. Furthermore, SIMMOL provides 3D visualization of the molecular structure with user-defined representation of relevant atoms, bonds, peptide planes, coordinate systems, and anisotropic interaction tensors. In this regard, we should note that SIMMOL is a greatly enhanced successor for the preliminar Octave (*18*) implemented PDB2SIMPSON program recently introduced for optimization of biological solid-state NMR experiments (*19*).

THE SIMMOL PROGRAM

The main idea behind the SIMMOL program is sketched in Fig. 1. Based on the atomic coordinates for a molecular structure (Fig. 1a), SIMMOL assigns elements such as peptide planes and typical anisotropic interaction tensors to user-specified parts of the structure leading to a specific visual representation along with a file containing spin system parameters in the Tcl format of a typical SIMPSON input file (Fig. 1b). Using this and a given pulse sequence (visualized using SIMDPS), SIMPSON enables straightforward calculation of the corresponding NMR spectrum or relevant coherence transfer efficiencies as included in a SIMPLOT representation in Fig. 1c. Provided that the aim is structure determination, rather than pulse sequence design/optimization, the simulated spectrum may be compared to a corresponding experimental spectrum using nonlinear minimization to obtain structural parameters (Fig. 1d). We note that this paper focuses primarily on the SIMMOL steps in the upper row of Fig. 1 and we restrict to give examples on the utility of this tool for optimization and simulation of 2D solid-state NMR experiments for rotating powders as well as uniaxially oriented samples. The comparison to experimental spectra may be conducted using tools available in the SIMPSON software package (*17*).

In full analogy to the SIMPSON software, SIMMOL contains a Tcl interpreter (*20*) which gives the user a high degree of flexibility to control all steps from the input of atomic coordinates to the delivery of SIMPSON spin system input files and 3D interactive graphics for optimum visualization of the molecular geometry along with the NMR specific interaction tensors. This is accomplished via a relatively simple Tcl input file, with the commands listed in Table 1. The basic features of these commands and examples on typical operations are given below, while we refer to the internet release (see below) for more a detailed description of all commands.

Availability, Portability, and Requirements

Prior to any description of the various features, it appears relevant to mention that the SIMMOL program is released at http://nmr.imsb.au.dk as open-source software which gives the user full access to the algorithms in the program and freedom to compile, correct, modify, and extend the program under the terms of the GNU General Public License (*23*). The source code can be compiled on any platform with a C++ compiler and a Tcl language interpreter (open-source software (*20*)). This ensures full operation with respect to establishing the `spinsys` part of the SIMPSON input file, while the open-source Unix-based Geomview (*24*) program is required for interpretation and interactive 3D visualization of the OOGL (object oriented graphics language) output files. Finally, we should mention that precompiled and self-contained binary executables are freely available for various Unix platforms (including Linux/i386).

Initialization

A first step in numerical evaluation of advanced solid-state NMR experiments, for applications on multiple-spin systems in polypeptides, is to read in atomic coordinates representing the structure elements to be investigated. This may be accomplished using the Tcl command

```
set m [mload filename]
```

loading a PDB file from the disk (the command in square brackets) into a descriptor `m` (using the `set` Tcl command). The PDB file may represent a real polypeptide structure obtained from the PDB database (*25*) (e.g., XRD, NMR, or cryoelectron microscopy data) or synthetic structures obtained by, e.g., molecular modeling (*26*). Alternatively, it is possible to use

```
set m [mmake nres phi psi omega]
```

to create an ideal poly-L-alanine peptide with *nres* residues and the torsion angles specified.

Prior to further operation, it may be relevant (optional) to specify the output files for the Geomview 3D visualization and the `spinsys` part of the SIMPSON input file. If proton coordinates are not available (as is typically the case for PDB files from XRD studies), it may be relevant to add amide and H^{α} protons to the molecular structure. Furthermore, it may be relevant to orient the molecule with the long axis along z potentially followed by

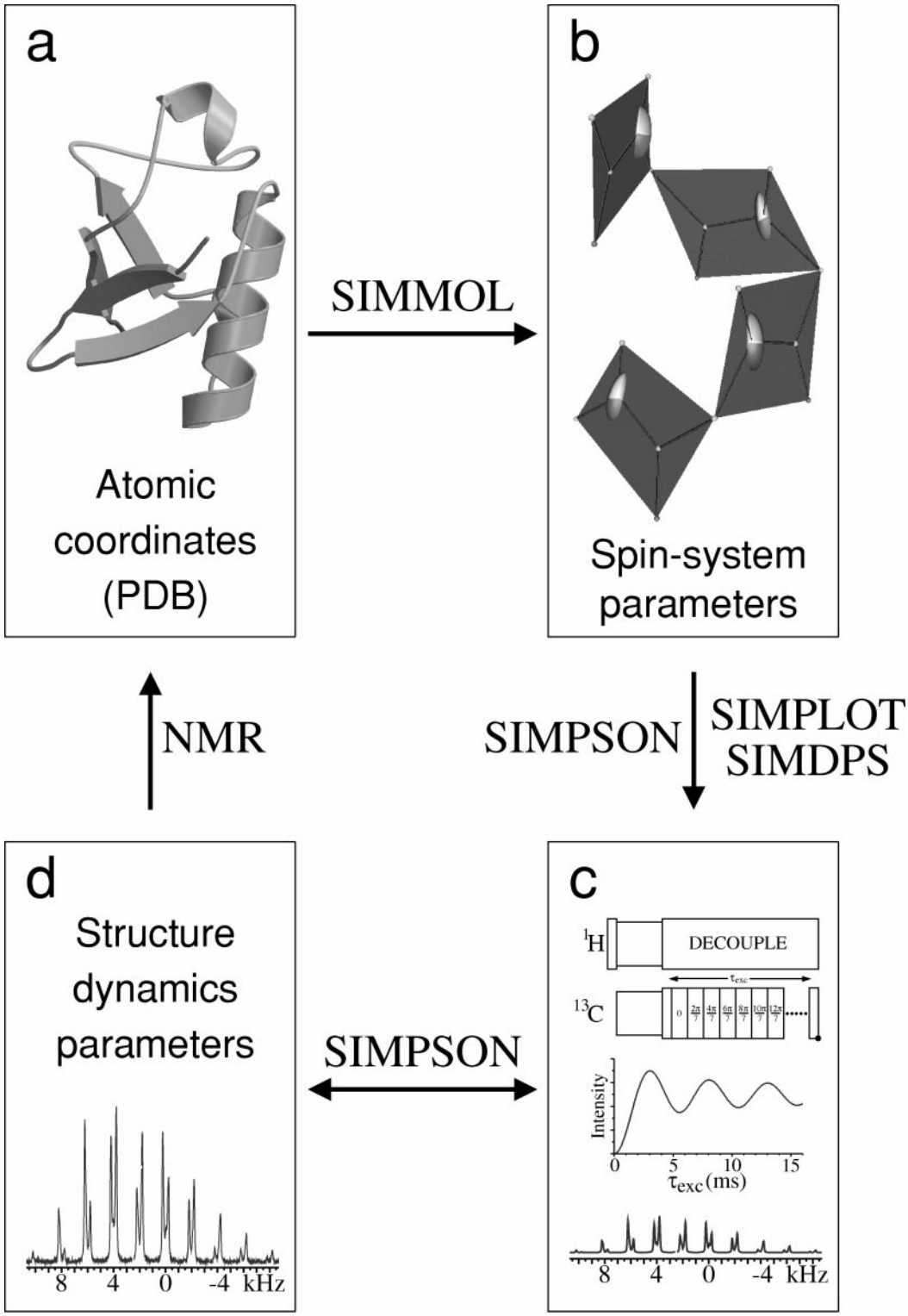

FIG. 1. Schematic representation of the combination of SIMMOL and SIMPSON for numerical simulation of multiple-pulse solid-state NMR experiments on polypeptide structures. (a) Atomic coordinates for a molecular structure in the PDB format, (b) typical SIMMOL interpretation of the structure and parameter output to the `spinsys` part of the SIMPSON input file, (c) SIMDPS and SIMPLOT visualization of a typical NMR pulse sequence, a coherence transfer function, and a spectrum simulated using SIMPSON, and (d) an experimental spectrum which through SIMPSON iterative fitting may provide NMR information about the molecular structure.

TABLE 1
Short Description of Tcl Commands for SIMMOL[a]

`mload` *filename*
Loads a PDB structure from disk and returns a descriptor.

`mmake` ϕ ψ *?ω? ?-geometry file?*
Creates a poly-L-alanine structure with the given torsion angles (ω defaults to 180°). The default geometry[b] may be overruled by a local geometry *file*. Returns a descriptor.

`munload` *desc*
Removes a structure from the memory and closes files.

`msave` *desc n filename*
Saves a PDB structure for the atoms selected in buffer *n*.

`msetspinsysfile` *desc filename ?-*[*sort*]*numbered?*
Establishes a file containing SIMPSON *spinsys* parameters generated using `mshift, mquadrupole, mjcoupling`, and `mdipole`.

`mclosespinsysfile` *desc ?-keep?*
Writes the selected spin system to the file specified using `msetspinsysfile` and clears the internal spin system selection unless the *-keep* flag is present.

`msetooglfile` *desc filename*
Creates a file containing OOGL commands for Geomview specifiers of the structure defined by the graphics commands.

`maddprotons` *desc -amide -alpha*
Adds coordinates for amide and H^{α} protons based on the C^{α}, N, and C′ coordinates. Returns added atom numbers.[c]

`mexchangeisotope` *desc n oldiso newiso*
Exchanges spin isotope from *oldiso* to *newiso* for atoms in buffer *n*. Returns affected atom numbers.[c]

`mzalign` *desc ?buffer?*
Orients the molecule with its long axis along *z* (when specified, only atoms in the buffer are used to orient the molecule). Returns the Euler angles for the rotation.[c]

`mrotate` *desc* {α β γ}
Rotates the molecule using the Euler angles α, β, and γ. Returns the rotation matrix.[c]

`mtranslate` *desc* {*x y z*}
Translates the molecule by the given vector.

`mtorsionangle` *desc (buffer|atom1 atom2 atom3 atom4)*
Prints out the backbone torsion angles for atoms in the buffer or prints the torsion angle for the four atoms.

`mloadtensors` *desc (filename|-default)*
Loads tensor values for anisotropic interactions (shift, quadrupole) from file *filename* or uses the default values in Table 2 (*-default*).

`mloadjcouplings` *desc (filename|-default)*
Loads one-bond *J* coupling constants from file *filename* or uses the default values in Table 2 (*-default*).

`mselect` *desc n type selection* . . .
Selects atom, residue, chain, or buffer *(type)* from buffer *n* as specified by PDB atom number, atom type, residue number, residue type *(selection)* with various logical variables. *type* and *selection* can be repeated. For each selected atom it returns a list containing atom number, atom type, residue number, residue type, and peptide-plane number.[c]

`mset` *desc specifiers*[d–h]
Sets global flags and values controlling the specifiers (tensors, colors, axes, etc.) and text output of the subsequent commands.

`mposition` *desc list*
Returns a list of atom coordinates for the given list (*list*) of buffers.

`mdistance` *desc atom1 atom2*
Returns the distance between atoms 1 and 2 in Å.

`msetcoordsys` *desc (a b c|matrix)*
Defines the current coordinate system based on the arguments. Returns the rotation matrix.[c]

`mdipole` *desc n n2 min max ?Specifiers?*[d, f]
specifiers of/parameters for the dipolar coupling between atoms in buffers *n* and *n2* with distances or couplings between *min* and *max*. Returns pairs of atom numbers, tensor magnitude, and orientation.

`mshift` *desc n ?specifiers?*[d,e, f]
Specifiers of/parameters for the chemical shielding tensor for atoms in buffer *n*. Returns atom number, tensor magnitude, and orientation.[c]

`mquadrupole` *desc n ?specifiers?*[d,e, f]
Specifiers of/parameters for the quadrupole coupling tensor for atoms in buffer *n*. Returns atom number, tensor magnitude, and orientation.[c]

`mjcoupling` *desc n n2 ?specifiers?*[d, f]
Specifiers of/parameters for the *J* coupling between atoms in buffers *n* and *n2*. Returns atom numbers, tensor magnitude, and orientation.[c]

`matom` *desc n ?specifiers?*[d]
Visualizes the atoms in buffer *n*. Returns atom numbers.[c]

`mbond` *desc n n2 ?specifiers?*[d, f]
Visualizes bonds between atoms in buffers *n* and *n2*. Returns pairs of atom numbers.[c]

`mbackbone` *desc n ?specifiers?*[d, f,h]
Visualizes the smoothed structure of the backbone atoms in buffer *n*. Returns atom numbers.[c]

`mplane` *desc n ?specifiers?*[d,g]
Visualizes peptide planes containing the C^{α} atoms in buffer *n*. Returns list of residues.[c]

`mline` *desc* {x_1 y_1 z_1} {x_2 y_2 z_2} *?specifiers?*[d, f]
Draws a line between the given coordinates.

`marrow` *desc* {x_1 y_1 z_1} {x_2 y_2 z_2} *?specifiers?*[d, f]
Draws an arrow from coordinates 1 to 2.

`mcylinder` *desc* {x_1 y_1 z_1} {x_2 y_2 z_2} *?specifiers?*[d]
Draws a cylinder between the given coordinates.

`mcone` *desc* {x_1 y_1 z_1} {x_2 y_2 z_2} *?specifiers?*[d]
Draws a cone from coordinates 1 to 2.

`msphere` *desc* {*x y z*} *?specifiers?*[d]
Draws a sphere at the given position.

`mpoly` *desc* {x_1 y_1 z_1}. . . {x_n y_n z_n} *?specifiers?*[d,g]
Shades the area between the given coordinates.

`mdingbat` *desc* {*x y z*} {α β γ} *dingbat ?specifiers?*[d, f]
Draws a dingbat (*text*, *dart arrow*, . . .) at the given position and orientation.

`mgeteulerangles` *matrix*
Returns Euler angles corresponding to a rotation matrix.

`mgeteulermatrix` {α β γ}
Returns the rotation matrix corresponding to the Euler angles.

`mmath` *a* (+| − | ∗ |/|*x*) *b*
Returns the result of the algebraic operation. *a* and *b* are numbers, vectors, or matrices.

[a] A more detailed description can be found in the SIMMOL package freely available at http://nmr.imsb.au.dk. Arguments enclosed by question marks are optional.

[b] N, C, and O bond lengths and angles: $r_{N–C^{\alpha}} = 1.458$ Å, $r_{C^{\alpha}–C'} = 1.525$ Å, $r_{C^{\alpha}–C^{\beta}} = 1.521$ Å, $r_{C'–N} = 1.329$ Å, $r_{C'–O} = 1.231$ Å, $\angle$(C′–N–C^{α}) = 121.7°, $\angle$(N–C^{α}–C′) = 111.2°, $\angle$(C^{α}–C′–N) = 116.2°, and $\angle$(N–C′–O) = 123.0° (*21*). $r_{N–H} = 1.07$ Å, $\angle$(C′–N–H) = 123°, $r_{C^{\alpha,\beta}–H^{\alpha,\beta}} = 1.09$ Å (*22*). All other angles are ideal tetrahedral angles of 109.4°.

[c] Return values depending on whether -[no]returnvalues is set.

[d] General visualization specifiers -size s, -color ({r g b}|cpk|hydropathy), -[no]nice.

[e] Tensor specifiers -[no]coordsys, -ellipsoid (unique|shielding)/-noellipsoid, -ellipsoidcut {*x y z*} /-noellipsoidcut, -[no]usecoordsys, -cutcolor {r g b}, -angles {α_{PE} β_{PE} γ_{PE}}, -magnitude {iso aniso eta}, -textsize s .

[f] Line specifiers -linewidth w, -segments s.

[g] Plane specifier -[no]solid.

[h] Backbone specifiers -helixcolor ({r g b} | hydropathy), -strandcolor ({r g b} | hydropathy), -turncolor ({r g b} | hydropathy).

a user-defined rotation of the molecule relative to the "laboratory fixed" coordinate system. All this may be accomplished by the commands

```
msetooglfile $m "test.oogl"
msetspinsystem $m "test.spinsys" -numbered
maddprotons $m -amide -alpha
mzalign $m
mrotate $m {90 90 0},
```

where the atoms are requested to be numbered successively in the `test.spinsys` file (optional) and the rotation of the molecule is conducted using the given Euler angles. We note that the molecule may be translated in space using the `mtranslate` command and that spin isotopes may be exchanged using `mexchangeisotope`.

Selections and Molecular Representation

The next step in a typical SIMMOL evaluation of the structure file is to select the relevant part of the molecular structure including the atoms, bonds, internuclear vectors, and peptide planes of interest. These elements form the framework for the subsequent attribution of anisotropic interaction tensors to the structure. The relevant parts of the structure may be selected using the command `mselect` typically used in constructions such as

```
mselect $m 1 residue 15..25 atom &!N,
```

specifying that all but the amide N atoms in residues 15 to 25 will be inserted into buffer 1 (&! represents "and not" among several logical variables operating on previously selected atoms). It should be noted that the `mselect` command offers the flexibility to include a repeated series of type (residue, atom, bond, helix, strand, turn, chain, plane, or buffer) and selection (residue numbers, atom type, residue type, helix number, peptide plane number, etc.) potentially preceded by logical modifiers (and: `&`, or: `|`, not: `!`, and all possible combinations of these) to specify any part of the molecule in one or several commands using the descriptors and buffers. For example, specific residues may be selected using the typical three-letter codes for the amino acids. The flexibility and straightforward combination with standard Tcl procedures may be exemplified by

```
mselect $m 1 atom C |CA |N
foreach residues {29 31..34 65} {
  mselect $m 2 residue $residues buffer &1
  ...
},
```

which in three distinct operations select the backbone ^{13}C and ^{15}N atoms in residues 29, 31–34, and 65 for further operation, and by

```
mselect $m 1 atom CA |C |O residue &4
mselect $m 2 atom CA |N |H residue &5 buffer |1,
```

which select the atoms defining the peptide plane. The latter two-step operation could also be accomplished using the operation `mselect $m 2 plane 4`. We should note that all logical operators are evaluated sequentially from left to right. Finally, it is relevant to mention that `mselect` (in full analogy to many of the other Tcl commands) returns information about the specific selections in the form of a list that for each atom contains information about the PDB file atom number, the atom type, residue number, residue type, and peptide-plane number, e.g., `{450 CA 29 LYS 28} {471 N 31 ILE 30}...`, which may be retrieved into a local variable using Tcl constructs of the type `set loclist [mselect ...]`.

Details on the output (graphics and return parameters) may be controlled globally using the `mset` command with a large number of optional arguments being specified as flags (preceded by -) and associated adjustables (e.g., numbers). For example, the general visualization specifiers allow definition of the relative size (`-size`), linewidth (`-linewidth`), color (`-color`), and quality of the graphical appearance `-[no]nice` for all types of graphical objects. The specifier `-[no]returnvalues` controls the return of parameters such as atom numbers, atom types, peptide plane numbers, Euler angles, rotation matrices, or anisotropic interaction parameters activated by subsequent commands. In addition to these comes a large number of specifiers for the graphical representation of tensors (e.g., `-[no]ellipsoid`, `-ellipsoidcut`, `-[no]coordsys`, `-magnitude`, `-angles`), peptide planes (`-[no]solid`), and special visualization of the backbone structure (`-helixcolor`, `-strandcolor`, `-turncolor`). A typical command would be

```
mset $m -noreturnvalues -color {0 0 1} -size 1.2,
```

where the arguments deactivate return of values from subsequent commands, specify the color blue (argument: {*red green blue*}, all three being in the range from 0 to 1), and scale geometric objects by a factor of 1.2 relative to default sizes. The arguments could also include, e.g., `-nice` to draw sphere and tensor surfaces with a high polygon resolution (also changeable within Geomview) and the arguments `cpk` or `hydropathy` to the `-color` specifier to activate the cpk color scheme or color labeling according to residue hydropathy (*27*). We note that all settings may be overruled locally by arguments to the commands activating the geometrical objects. A listing of available options can be found in the footnotes to Table 1.

With appropriate settings, atoms, bonds, and peptide planes may be visualized and parameterized using the commands

```
matom $m 1
mplane $m 1 -color {0 1 0}
mbond $m 1 2,
```

where atoms (balls) and peptide planes (spanned by C^α and H^N in the selected residue along with the carbonyl oxygen and C^α in the preceding residue) are attributed to the atoms in buffer 1. Likewise, covalent bonds are established between atoms in buffers 1 and 2 (defined using `mselect`).

In addition to these fundamental commands, SIMMOL has numerous other commands for custommade visualization, e.g., in the form of lines, arrows, cylinders, spheres, polygons, and various dingbats as described briefly in Table 1. For example, the polypeptide backbone in buffer 1 may be given in a smoothed representation using

```
mbackbone $m 1 -color hydropathy -linewidth 5
```

with specification of hydropathy color labeling and a certain linewidth.

An important feature of SIMMOL is the possibility for establishing local coordinate systems that may subsequently be used to define the orientation of an NMR interaction tensor. The coordinate system is defined from the atomic coordinates of three atoms

```
set pos [mposition $m {1 2 3}]
msetcoordsys $m [lindex $pos 0][lindex $pos 1] \
               [lindex $pos 2],
```

where the buffers 1, 2, and 3 contain the three involved atoms *A*1, *A*2, and *A*3 selected a priori using commands such as `mselect $m 1 atom N residue &6`. The newly established coordinate system has its *x* axis oriented along $\mathbf{A1}-\mathbf{A2}$, *z* along $x\times(\mathbf{A3}-\mathbf{A2})$, and *y* perpendicular to *x* and *z*, where **Ai** denotes the vector from origin to the position of atom *Ai*. In combination with algebraic operations using the `mmath` command, this feature proves useful for the establishment of, e.g., anisotropic interaction tensors in the peptide side chains for which default parameters may not be available (*vide infra*). Finally, the `mdingbat` command allows the association of text and special markers to emphasize important elements in the molecular structure.

Tensor Representation

While the atoms, bonds, and peptide planes may be controlled exclusively on basis of the atomic coordinates, the most important part of the SIMMOL operation, namely the anisotropic nuclear spin interaction tensors, requires further definition in terms of the nuclear spin Hamiltonian. In general, simulations are conducted on basis of a high-field truncated Hamiltonian, where the internal part of a first-order interaction may take the typical form

$$H_\lambda = \sum_{m=-2}^{2} \omega_{\lambda,0}^{(m)} e^{im\omega_r t}\mathcal{O}_\lambda, \qquad [1]$$

with $\omega_r/2\pi$ denoting the spinning frequency and $\mathcal{O}_\lambda$ the spin operator. The Fourier coefficients may be written as

$$\omega_{\lambda,0}^{(m)} = \omega_{\text{iso}}^{\lambda}\delta_{m,0} + \omega_{\text{aniso}}^{\lambda}\left\{D_{0,-m}^{(2)}\left(\Omega_{\text{PR}}^{\lambda}\right) - \frac{\eta^\lambda}{\sqrt{6}}\left[D_{-2,-m}^{(2)}\left(\Omega_{\text{PR}}^{\lambda}\right) + D_{2,-m}^{(2)}\left(\Omega_{\text{PR}}^{\lambda}\right)\right]\right\} d_{-m,0}^{(2)}(\beta_{\text{RL}}), \qquad [2]$$

using the angular frequencies $\omega_{\text{iso}}^{\text{CS}} = \omega_0\delta_{\text{iso}}$, $\omega_{\text{aniso}}^{\text{CS}} = \omega_0\delta_{\text{aniso}}$, $\omega_{\text{iso}}^{J} = -2\pi\sqrt{3}J_{\text{iso}}$, $\omega_{\text{iso}}^{D} = 0$, $\omega_{\text{aniso}}^{D} = \sqrt{6}b_{\text{IS}}$, $\omega_{\text{iso}}^{Q} = 0$, and $\omega_{\text{aniso}}^{Q} = 2\pi\sqrt{6}C_Q/(4I(2I-1))$ for the chemical shift, scalar *J* coupling, dipolar coupling, and quadrupolar coupling interactions, respectively. A more detailed description can be found in Ref. (*17*).

The orientation of an anisotropic tensor is expressed using second-rank Wigner ($D^{(2)}$) and reduced Wigner ($d^{(2)}$) matrices describing coordinate transformations from the principal-axis frame (P^λ) to the laboratory-fixed frame (L). For the present purpose it proves convenient to let the transformations further involve a peptide plane frame (E), a molecule (or crystal) fixed frame (C), and a rotor-fixed frame (R). With $\Omega_{XY}^{\lambda} = \{\alpha_{XY}^{\lambda}, \beta_{XY}^{\lambda}, \gamma_{XY}^{\lambda}\}$ denoting the Euler angles relating two frames *X* and *Y*, the transformations relating *P* and *R* may be written

$$D_{m',m}^{(2)}\left(\Omega_{PR}^{\lambda}\right) = \sum_{m'',m'''=-2}^{2} D_{m',m''}^{(2)}\left(\Omega_{PE}^{\lambda}\right) \times D_{m'',m'''}^{(2)}\left(\Omega_{EC}^{\lambda}\right) D_{m''',m}^{(2)}(\Omega_{CR}), \qquad [3]$$

while *R* is related to *L* by a Wigner rotation using $\alpha_{RL} = \omega_r t$, $\beta_{RL} = \tan^{-1}\sqrt{2}$, and $\gamma_{RL} = 0$ for a sample spinning at the magic angle while $\Omega_{RL} = (0, 0, 0)$ for a static sample. With Ω_{CR} describing the orientation of the individual crystallite relative to *R* (the "powder angles"), we are left with the need for specification of the most relevant transformations,

$$P^\lambda \xrightarrow{\Omega_{PE}^{\lambda}} E \xrightarrow{\Omega_{EC}} C, \qquad [4]$$

relating the principal axis frame of the anisotropic tensors to the crystal/molecular frame *C*. It should be emphasized that the rotations in Eq. [3] represent rotations of the coordinate system of reference, while commands such as `mrotate` and `mzalign` operate in the opposite manner as they rotate the object (i.e., the molecule) relative to the reference frame.

TABLE 2

Typical Magnitudes and Orientations of Chemical Shift, Scalar *J* Coupling, Dipole–Dipole Coupling, and Quadrupolar Coupling Tensors within the Amino Acid Residue or Peptide Plane of Polypeptides[a]

Chemical shift	Spin	δ_{iso}^{CS}	δ_{aniso}^{CS}	η^{CS}	α_{PE}^{CS}	β_{PE}^{CS}	γ_{PE}^{CS}	Ref.
	$^{1}H^{N}$	9.3	7.7	0.65	90	−90	90	*31*
	$^{13}C^{\alpha}$[b]	50	−20	0.43	90	90	0	*32, 33*
	$^{13}C'$	170	−76	0.90	0	0	94	*34–36*
	^{15}N	119	99	0.19	−90	−90	−17	*31, 35–37*
J and dipolar coupling	Spins	J_{iso}^{IS}[c]	$b_{IS}/2\pi$[d]	r_{IS}^{e}	β_{PE}^{D}	γ_{PE}^{D}		
	$^{1}H^{\alpha}$–$^{13}C^{\alpha}$	140	−23328	1.090	—[f]	—[f]		
	$^{1}H^{N}$–^{15}N	−92	11341[g]	1.024[g]	90	0		
	$^{13}C^{\alpha}$–$^{13}C'$	55	−2142	1.525	90	120.8		
	$^{13}C^{\alpha}$–$^{13}C^{\beta}$	35	−2159	1.521	—[f]	—[f]		
	$^{13}C^{\alpha}$–^{15}N	−11	988	1.458	90	115.3		
	$^{13}C'$–^{15}N	−15	1305	1.329	90	57		
Quadrupolar coupling	Spin		C_Q	η^{Q}	α_{PE}^{Q}	β_{PE}^{Q}	γ_{PE}^{Q}	Ref.
	$^{2}H^{N}$[h]		0.210	0.15	−90	90	0	*38*
	$^{2}H^{\alpha}$		0.168	0.10	—[f]	—[f]	—[f]	*39, 40*
	^{14}N		3.21	0.32	0	0	0	*35, 41*
	^{17}O[h,i]		8.3	0.28	−90	90	0	*42*

[a] The Euler angles relate the principal axes frames (P^{λ}) to the peptide plane (E) having x_E along N–H and z_E being the normal to the plane (cf. Fig. 2a). Chemical shifts (δ_{iso}^{CS}, δ_{aniso}^{CS}; defined in Ref. *17*) are in ppm relative to TMS (^{1}H, ^{13}C) and liq. NH_3 (^{15}N). Scalar J and dipolar couplings (J_{iso}^{IS} and $b_{IS}/2\pi$) are given in Hz, the internuclear distance r_{IS} in Å, and the quadrupolar coupling in MHz. Due to axial symmetry the dipolar coupling α_{PE} (and α_{PC}) angle can be chosen arbitrarily.

[b] The orientation of the $^{13}C^{\alpha}$ chemical shielding tensor depends on the secondary structure and may vary significantly from the given angles (*43*).

[c] Scalar coupling constants taken from Ref. (*44*).

[d] Dipole couplings calculated from r_{IS}.

[e] Typical N–H and C–H bond lengths are taken from Ref. (*22*), while N–C and C–C bond lengths are taken from Ref. (*21*). We note that SIMMOL automatically calculates the bond length; dipolar coupling constants, and dipolar coupling Ω_{PC}^{D} Euler angles directly from the PDB structure without reference to the tabelized values.

[f] The Euler angles depend on the secondary structure.

[g] A somewhat lower value of $b_{IS}/2\pi = 9.9$ kHz (corresponding to $r_{IS} = 1.07$ Å) is typically used for peptides oriented in uniaxially aligned phospholipid bilayers (*45*).

[h] We note that the magnitude and in particular the orientation of these tensors may be influenzed by hydrogen bonding.

[i] We assume that the ^{17}O quadrupolar coupling tensor is oriented with its unique element Q_{zz} along the C′–O bond axis and Q_{yy} perpendicular to the peptide plane.

For polypeptides it is possible to establish quite reasonable parameters for the chemical shift, scalar J coupling, and quadrupolar coupling tensors using typical (or average) values reported for a large number of amino acids and small peptides. In most cases the parameters describing both the magnitude and the orientation of the anisotropic parts of these tensors relative to the peptide plane (i.e., Ω_{PE}) exhibit only relatively small dependency on the residue type and the local structure (*28–30*). While these minor variations may be very important probes for structure determination, the variation is sufficiently small that the typical values in almost any case will be sufficiently precise for simulation and optimization of pulse sequences for application on real structures. The dipolar coupling tensors are even simpler to establish in the sense that their magnitudes are related to the internuclear distance r_{ij} as $b_{ij} = -\gamma_i\gamma_j\mu_0\hbar/(r_{ij}^3 4\pi)$, the tensor is axially symmetric, and the unique element is oriented along the internuclear axis. Thus, for the dipolar coupling SIMMOL establishes the $P^{\lambda} \rightarrow C$ Euler angles Ω_{PC}^{λ} without the intermediate $P^{\lambda} \rightarrow E$ coordinate transformation.

Overall, the various tensors within the peptide plane (or a single residue) may be described by the typical parameters given in Table 2 with the principal axis frame to peptide plane Euler angles Ω_{PE}^{λ} relating the tensors to a peptide plane coordinate system E with the x axis along the N–H bonding and the z axis perpendicular to the peptide plane as visualized in Fig. 2a. These or user-defined parameters for the magnitude and orientation of relevant interaction tensors may be loaded into SIMMOL using

```
mloadtensors "tensor.in"
mloadjcouplings "jcouplings.in",
```

which loads user-specific values from the specified files, while an argument `-default` would cause loading of the default chemical shift, quadrupolar, and J coupling parameters in Table 2.

Equipped with the atomic coordinates for the molecular structure and the parameters for the peptide plane interaction tensors, SIMMOL allows straightforward association of the NMR

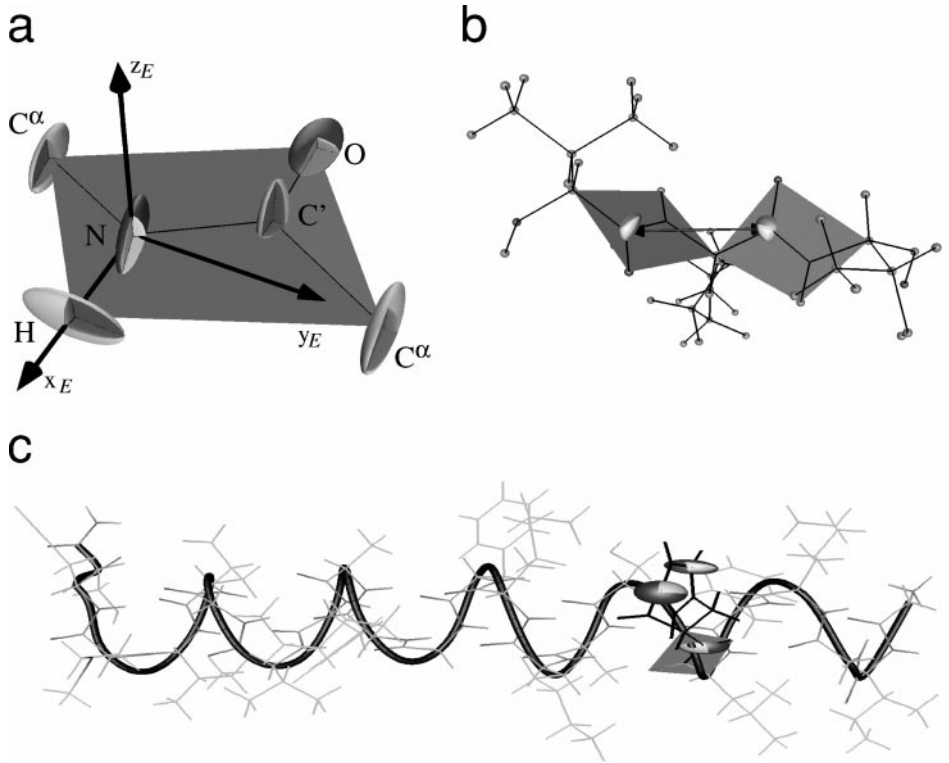

FIG. 2. Geomview interpreted 3D graphical output from the SIMMOL program. (a) ^{1}H, ^{13}C, and ^{15}N chemical shift and ^{17}O quadrupolar coupling tensors within a peptide plane (the x_E, y_E, z_E axes define the peptide plane coordinate system E). (b) Visualization of a pair of peptide planes in ubiquitin (PDB ID: 1D3Z) with different chemical shift and dipolar coupling (i.e., internuclear axes) tensors specified. (c) Demonstration of various graphical features, including the `mbackbone` command, for magainin (PDB ID: 2MAG).

relevant tensor information to the molecular graphics and export of the corresponding parameters to the `spinsys` part of the SIMPSON input file. This may be accomplished using a set of simple commands giving a high degree of flexibility. Consider, for example, the commands

```
mselect $m 1 atom N
mshift $m 1 -coordsys -ellipsoid unique \
            -ellipsoidcut {1 0 0} \
            -color {1 0 1} -nice -noreturnvalues,
```

which select all amide nitrogens from the molecular structure into buffer 1 and associate anisotropic chemical shielding tensors to these using the `mshift` command. For the purpose of illustration we included a few specifiers in the latter command, among which the first three introduce the N-spin principal axis coordinate system and a tensor ellipsoid in a "unique" element representation (default). This gives a good visual discrimination of the three individual principal axis components with weighting according to uniqueness, obtained using relative axis lengths of $1:0.5:0.25$ for the δ_{zz}, δ_{xx}, δ_{yy} elements (ordered as $|\delta_{zz}-\delta_{\rm iso}| \geq |\delta_{xx}-\delta_{\rm iso}| \geq |\delta_{yy}-\delta_{\rm iso}|$ with $\delta_{\rm iso} = \frac{1}{3}(\delta_{xx}+\delta_{yy}+\delta_{zz})$, which emphasizes the unique tensor element δ_{zz}. We note that all coodinate transformations from the principal axis frame initially rotate around the axis corresponding to the the unique element in accord with common practice (*46*). Alternatively, the tensor may be visualized in the more conventional shielding representation with the longest axis denoting the most shielded δ_{33} element using `shielding` as argument to the `-ellipsoid` specifier. In this case the relative axis lengths are obtained as $l_{ii} = \delta_{11} - \delta_{ii} + (\delta_{11} - \delta_{33})/3$ using the standard ordering $\delta_{33} \leq \delta_{22} \leq \delta_{11}$ (corresponding to $\sigma_{33} \geq \sigma_{22} \geq \sigma_{11}$ on the shielding scale) (*47*). The magnitude and orientation of the tensor relative to the peptide plane and molecule are highlighted by a missing octant cut out of the tensor ellipsoid using the `ellipsoidcut` specifier with the vector argument specifying which octant.

In the above `mshift` example, it is assumed that the parameters for the tensor take default or user-specific values loaded using `mloadtensors`. Alternatively, they may be entered or overruled locally by attaching the `-magnitude` and `-angles` specifiers to the `mshift` commandline. In addition to providing output to the OOGL file for interactive Geomview 3D visualization, the `mshift` command writes the tensor parameters to a file containing the `spinsys` part of the SIMPSON input file. We note that the Tcl `set` command may be employed to retrieve the spin system parameters for internal use, e.g., using

```
set spar [mshift $m 1 -returnvalues],
```

which for the involved atoms will attribute a list containing the atom number, the tensor magnitude ($\delta_{\text{iso}}^{\text{CS}}$, $\delta_{\text{aniso}}^{\text{CS}}$, and η^{CS}), and the tensor orientation (Ω_{PC}) to the local variable `spar`. The individual values may be extracted using the Tcl `[lindex $spar i]` construct which returns the ith element from the list `spar` (starting with $i = 0$).

Spin system parameters and graphical representation of dipole–dipole coupling tensors involving, e.g., C^{α} and C' carbons separated by 2.5–4 Å, may be established using

```
mselect $m 1 atom CA |C
mdipole $m 1 1 2.5AA 4.0AA,
```

which select all C^{α} and C' carbons in buffer 1 and add an internuclear vector between those satisfying the indicated distance selections in the molecular structure. Furthermore, `mdipole` provides all relevant information about the dipolar coupling constants and Ω_{PC} orientational angles required for numerical simulations. Likewise, quadrupolar coupling and scalar J coupling interactions may be accessed using the `mquadrupole` and `mjcoupling` commands, respectively.

Execution

The SIMMOL program may be executed on a standard Unix (or Linux) environment using the command

```
>simmol struct.mol,
```

where > is the Unix prompt and `struct.mol` is the user-specific input file. The output (`vide infra`) will typically be the files `struct.oogl` and `struct.spinsys` containing commands for Geomview 3D interactive graphics and a spin system file which may be inserted as part of a typical SIMPSON input file (or used in combination with other simulation software), respectively.

Cooperation between SIMMOL and SIMPSON can be performed using the Tcl `exec` command executing an external program from within the Tcl code. This may be accomplished by the commands

```
msetspinsysfile $m "struct.spinsys" -numbered
...
mclosespinsysfile $m
exec simpson psg.in,
```

where `psg.in` is a normal SIMPSON input file (*17*) which instead of the normal `spinsys` section contains the line

```
source "struct.spinsys".
```

Graphical and Parameter Output

By appropriate combination of the molecular structure and NMR interaction tensor commands described above, SIMMOL offers the flexibility to establish reliable values for essentially all relevant tensor parameters and visualization of relevant parts of the molecular structure for optimum description of the results from numerical simulations aimed at structure determination or optimization of experimental procedures. This is demonstrated in Figs. 2 and 3 illustrating typical graphical output (visualized using Geomview) and parameter input/output files, respectively, from the SIMMOL program.

Figure 2a gives a SIMMOL representation of the various ^{1}H, ^{13}C, and ^{15}N chemical shift and ^{17}O quadrupolar coupling tensors relative to the peptide plane corresponding to the parameters in Table 2. To give a first demonstration of the flexibility of SIMMOL as a tool to visualize relevant parts of a molecular structure, Figs. 2b and 2c show the Val_6-Lys_7 residues of the ubiquitin 1D3Z liquid-state NMR structure (*49, 50*) and a full representation of the magainin 2MAG micelle liquid-state NMR structure (*51*), respectively, with different peptide planes, tensors, and bondings highlighted using various of the commands described above. It should be noted that the commands for color labeling of the structure and tensor elements may not be fully appreciated in the grayscale representation in Fig. 2.

The straightforward and highly flexible control of the molecular graphics and the parameter output for numerical simulations is demonstrated in Fig. 3 showing the Tcl input file and the resulting SIMPSON `spinsys` file corresponding to the graphics in Fig. 2c. The SIMMOL Tcl input file in Fig. 3a uses a large number of the commands described above to load the atomic coordinates, specify output files, orient the molecule, and select the relevant residues, backbone, atoms, peptide planes, bonds, and interaction tensors with specific attention to the ^{15}N chemical shift tensors for the histidine residue (*48*). Using this file as input to SIMMOL leads to the 3D graphics in Fig. 2c and the Tcl SIMPSON `spinsys` file in Fig. 3b. The latter file, which specifies the relevant RF channels, nuclei, and nuclear spin interaction parameters, can be used directly as part of a standard SIMPSON input file which additionally defines the actual solid-state NMR experiment.

The details of the anisotropic interactions may be very important for the setup of reliable numerical simulations of pulse sequences operating on specific structures associated with characteristic nuclear spin interaction tensors. While the magnitude of the various interaction tensors may readily be inserted into the SIMPSON input file, it is often much more tedious and errorprone to manually establish the orientation of the relevant anisotropic tensors. These difficulties, which in practice may severely hamper appropriate evaluation of pulse sequences on relevant structures, may be solved using SIMMOL as described above. Seen in this perspective, SIMMOL may be considered a "sample changer" to the SIMPSON "computer spectrometer" and may as so be useful for pulse sequence evaluations and the setup of initial parameters for iterative fitting of experimental spectra. In this regard, we should emphasize the importance of having SIMMOL combining molecular graphics and specific information about the anisotropic tensors for numerical simulations. This enables direct verification of the tensor orientations in the frame of the molecular structure otherwise

a

```
set m [mload "2MAG.pdb"]                                          (1)
msetooglfile $m "2MAG.oogl"
msetspinsysfile $m "2MAG.spinsys" -numbered
mloadtensors $m -default

mselect $m 1 atom *                                               (2)
mbackbone $m 1 -linewidth 15 -nice -color {0 0 0}

mselect $m 2 residue !HIS
mbond $m 2 2 -linewidth 4

mselect $m 2 residue HIS
mbond $m 1 2 -linewidth 6 -color {0 0 0}
mplane $m 2 -solid -size 1.1
mselect $m 2 atom &N
mshift $m 2 -size 2 -nice

proc setcoordsys {m c1 n c2} {                                    (3)
  mselect $m 8 atom $c1 residue &HIS
  mselect $m 2 atom $n residue &HIS
  mselect $m 9 atom $c2 residue &HIS
  set pos [mposition $m {8 2 9}]
  set c1pos [lindex $pos 0]
  set npos [lindex $pos 1]
  set c2pos [lindex $pos 2]
  set cnc [mmath [mmath $c1pos + $c2pos] - $npos]
  msetcoordsys $m $cnc $npos $c1pos
}

mset $m -usecoordsys -size 2 -nice                                (4)
setcoordsys $m CG ND1 CE1
mshift $m 2 -angles {14.2 0 0} \
            -magnitude {247p -213p 0.37}
setcoordsys $m CD2 NE2 CE1
mshift $m 2 -angles {13.6 0.7 0} \
            -magnitude {168p -99p 0.74}

mselect $m 1 atom N* residue &HIS                                 (5)
mselect $m 2 atom H*
set NH [mdipole $m 1 2 0AA 1000Hz -linewidth 0]

mselect $m 1 atom !*
foreach H $NH {
  mselect $m 1 atom |[lindex [lindex $H 0] 1]
}
mdipole $m 1 1 0AA 0Hz -linewidth 0

munload $m
```

b

```
spinsys {
#      1       2       3       4       5       6       7
#    97N 103ND1 106NE2   107H   108HA 111HD1 113HE1
#
 channels 15N 1H
 nuclei   15N 15N 15N 1H 1H 1H 1H
 shift 1 119.3p 97.7p 0.13 -139.44 107.93 178.53
 shift 2 247p -213p 0.37 -38.458 54.222 -175.41
 shift 3 168p -99p 0.74 -152.6 125.07 4.3626
 dipole 1 4 12176 0 85.23 -12.473
 dipole 1 5 1399 0 144.49 170.9
 dipole 2 6 12946 0 60.326 -61.223
 dipole 2 7 1232.7 0 137.15 -136.35
 dipole 3 7 1231 0 91.849 -86.753
 dipole 4 5 -5422.7 0 128.74 169.37
 dipole 4 6 -894.27 0 124.35 -109.71
 dipole 4 7 -331.42 0 133.95 -124.29
 dipole 5 6 -1245.2 0 104.29 -80.589
 dipole 5 7 -654.21 0 124.15 -99.04
 dipole 6 7 -7328.6 0 144.12 -169.9
}
```

FIG. 3. Tcl (a) input file corresponding to the SIMMOL graphics of magainin in Fig. 2c. Labels: (1) Initialization, load PDB file, set output files, and load tensors. (2) Draw backbone and bonds emphasizing the His residue. Draw peptide plane and ^{15}N amide shift tensor. (3) Tcl routine to set coordinate system on histidine side chain nitrogens (using msetcoordsys) with its x-axis along the C–N–C interception and z perpendicular to the ring. (4) Insert the shift tensors according to Wei *et al.* (*48*). (5) Find all hydrogens coupled to the three nitrogens with couplings larger than 1200 Hz and find their mutual dipole couplings. (b) Output file (for the `spinsys` part of a SIMPSON input file) generated by the above input file and corresponding to the SIMMOL graphics of magainin in Fig. 2c.

at best being limited to the individual peptide plane. Furthermore, the visual representation proves very useful for selection of the appropriate spin systems from peptide structures using different isotope labeling schemes. Finally, the graphics may be very useful for precise documentation (e.g., in publications) of anisotropic tensor information in relation to molecular structure. This applies obviously to solid-state NMR, but may also be useful in relation to liquid-state NMR using residual anisotropies for structure refinement (*52, 53*).

NUMERICAL SIMULATIONS USING SIMMOL IN COMBINATION WITH SIMPSON

With the SIMMOL software established, it is relevant to demonstrate its use in combination with SIMPSON for numerical simulation of biological solid-state NMR experiments. Considering that biological solid-state NMR during the past decade to a large extent has been based on two different experimental strategies with one relying on rotating powder samples (*54*) and the other relying on peptides incorporated into uniaxially oriented phospholipid bilayers (*55*), we will demonstrate typical applications for numerical simulation and optimization of solid-state NMR spectra within these two directions.

2D MAS N(CO)CA, N(CA)CB, and N(CA)CX Experiments for Spectral Assignment

In order to extract structural information from solid-state NMR spectra of uniformly ^{13}C,^{15}N-labeled peptides, it is necessary to establish connectivity information which allows assignment of all ^{13}C and ^{15}N resonances to their amino acids. This may be accomplished using a variety of different experiments as recently described by Tycko (*56*), Straus *et al.* (*57*), Hong and Colleagues (*58, 59*), McDermott *et al.* (*60*), Bak *et al.* (*19*), Pauli *et al.* (*61*), and Rienstra *et al.* (*62*). The method of choice obviously depends on the desired resonance correlations but also on the actual experimental conditions (e.g., sample spinning speed and available RF power). Furthermore, it may depend on the secondary structure influencing the dipolar couplings through which through-space correlations must be established. In all cases, it is fundamental to use the method which in the most efficient and selective manner establishes the correlations. This ensures the maximum sensitivity, the best spectral resolution, and reduces the risk for ambiguities in the resonance assignments. With these aspects in mind, we recently compared the performance of various pulse sequence building blocks to establish efficient pulse sequences for interresidue N_i–C^{α}_{i-1} and intraresidue N_i–C^{β}_i correlations (*19*). This led to the conclusion that N(CO)CA and N(CA)CB correlations can be obtained with very high sensitivity and selectivity using the pulse sequence in Fig. 4a based on DCP (*63*) for ^{15}N to ^{13}C transfer and C7 (*64*) for ^{13}C to ^{13}C transfer with appropriate parameters for the spinning speed, RF field strength, and carrier offsets. It was demonstrated that more than 50% efficiency can be obtained for

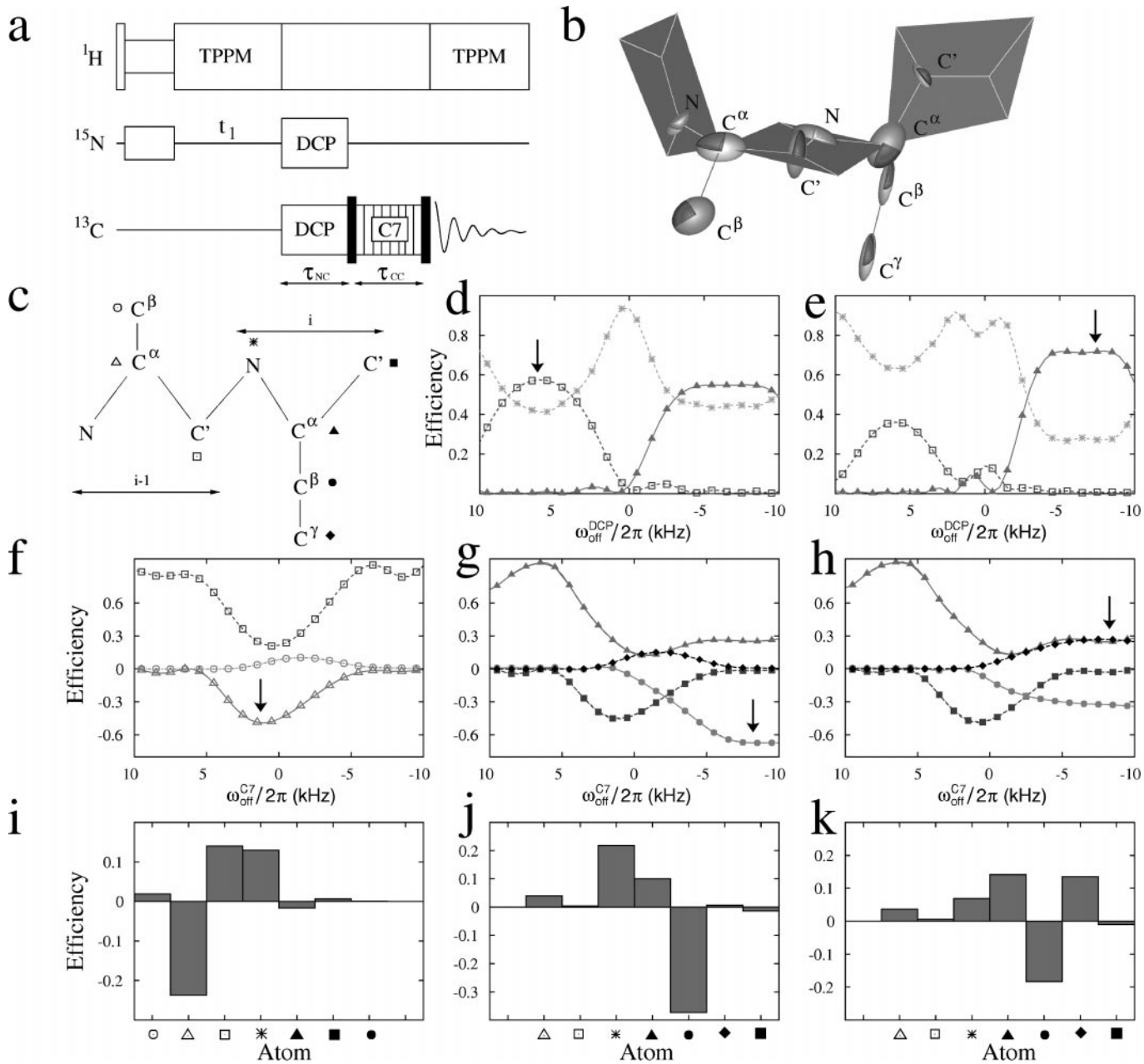

FIG. 4. Numerical simulations of DCP- and C7-based N(CO)CA, N(CA)CB, and N(CA)CX experiments on uniformly ^{15}N,^{13}C-labeled peptides obtained using SIMMOL in combination with SIMPSON. The pulse sequence in (a) employs $\omega_r/2\pi = 5$ kHz, $\omega_{RF}^{DCP,C}/2\pi = 35$ kHz, $\omega_{RF}^{DCP,N}/2\pi = 30$ kHz, and $\omega_{RF}^{C7}/2\pi = 35$ kHz. (b) SIMMOL representation of the Gln$_{40}$-Gln$_{41}$ fragment of ubiquitin (PDB ID: 1D3Z) with indication of peptide planes and tensors used for the numerical simulations. (c) Same fragment as in (b) but with each atom labeled by a symbol used in the following graphs. (d–h) Coherence transfer efficiencies for (d, e) heteronuclear DCP for optimum transfer to C′ (d) and C$^\alpha$ (e) based on the $\{^{13}C'_{33}$–$^{15}N_{34}$–$^{13}C^\alpha_{34}\}$ three-spin system and for (f–h) homonuclear C7 transfer from (f) $^{13}C'_{33}$ in the $\{^{13}C'_{33}$–$^{13}C^\alpha_{33}$–$^{13}C^\beta_{33}\}$ spin system (i.e., N(CO)CA), (g) $^{13}C^\alpha_{32}$ in the $\{^{13}C'_{32}$–$^{13}C^\alpha_{32}$–$^{13}C^\beta_{32}$–$^{13}C^\gamma_{32}\}$ spin system (i.e., N(CA)CB), and (h) $^{13}C^\alpha_{34}$ in the $\{^{13}C'_{34}$–$^{13}C^\alpha_{34}$–$^{13}C^\beta_{34}$–$^{13}C^\gamma_{34}\}$ spin system (i.e., N(CA)CX). For each graph the optimum condition is indicated by a vertical arrow. (i–k) Transfer efficiencies for seven of the eight spins of the type in (b) obtained using the pulse sequence in (a) with (i) N(CO)CA: $\tau_{DCP} = \tau_{C7} = 1.2$ ms, $\omega_{off}^{DCP}/2\pi = 6.0$ kHz and $\omega_{off}^{C7}/2\pi = 1.5$ kHz, (j) N(CA)CB, and (k) N(CA)CX: $\tau_{DCP} = 1.8$ ms, $\tau_{C7} = 1.2$ ms, $\omega_{off}^{DCP}/2\pi = -7.5$ kHz and $\omega_{off}^{C7}/2\pi = -8.5$ kHz (see text).

each of the relevant transfers and less than 10% efficiency for undesired transfer to the third spin in typical $^{13}C'$–^{15}N–$^{13}C^{\alpha}$ and $^{13}C'$–$^{13}C^{\alpha}$–$^{13}C^{\beta}$ three-spin systems. The high transfer efficiencies, which compare (*19*) very favorably with those provided by many commonly used pulse sequence elements, are ascribed to the fact that both sequence elements are γ-encoded (*65*) and relatively selective in their coherence transfers.

Using different physical conditions and interactions tensors for various three-, four-, and seven-spin systems involving $^{13}C_{i-1}^{\beta}$, $^{13}C_{i-1}^{\alpha}$, $^{13}C'_{i-1}$, $^{15}N_i$, $^{13}C_i^{\alpha}$, $^{13}C_i^{\beta}$, $^{13}C_i^{\gamma}$, and $^{13}C'_i$ from the Gln_{31}, Asp_{32}, Lys_{33}, and Glu_{34} residues of ubiquitin as exemplified graphically by the Gln_{40}-Gln_{41} (a β-sheet area in ubiquitin that is better visualized) residues in Fig. 4b, we demonstrate here typical elements of a pulse sequence optimization using SIMPSON in combination with SIMMOL. We take the DCP-C7 N(CO)CA, N(CA)CB, and N(CA)CX 2D pulse sequence in Fig. 4a as a specific example and refer to previous work for comparison with pulse sequences using different building blocks (*19*). The present physical conditions involve MAS with a spinning frequency $\omega_r/2\pi = 5$ kHz and a maximum of 35-kHz RF field strength on the ^{13}C and ^{15}N RF channels. This should ensure sufficient room for efficient ^{1}H decoupling (not shown) with an RF field strength exceeding that on the low-γ channels by the recommended factor of at least 3 (*66*) using commonly available solid-state NMR equipment.

On basis of the central $^{13}C'_{33}$–$^{15}N_{34}$–$^{13}C^{\alpha}_{34}$ three-spin system (see symbol labeling in Fig. 4c), we optimize the heteronuclear DCP transfers to 58 and 72% efficiency for the N_i to C'_{i-1} and N_i to C_i^{α} transfers, respectively, by variation of the ^{13}C RF carrier frequency ($\omega_{\text{off}}^{\text{DCP}}/2\pi$) and the mixing time τ_{DCP} to achieve optimum efficiency and selectivity as illustrated graphically in Figs. 4d and 4e. These curves reveal that a DCP pulse sequence with $\omega_{\text{RF}}^{\text{DCP,C}} = \omega_{\text{RF}}^{\text{DCP,N}} + \omega_r/2\pi = 35$ kHz with on-resonance irradiation at the desired carbon and appropriate mixing time (N_i–C'_{i-1}: $\omega_{\text{off}}^{\text{DCP}}/2\pi = 6$ kHz and $\tau_{\text{DCP}} = 6\tau_r$; N_i–C_i^{α}: $\omega_{\text{off}}^{\text{DCP}}/2\pi = -7.5$ kHz and $\tau_{\text{DCP}} = 9\tau_r$ with $\tau_r = 2\pi/\omega_r$) provides transfer efficiencies close to the theoretical maximum of γ-encoded recoupling sequences (*65*) and that the transfer can be made highly selective using DCP (less than 1% loss to the undesired carbon for the two transfers). Both features are fundamental for the overall performance of the 2D dimensional experiment additionally relying on a subsequent homonuclear transfer step.

For the homonuclear transfers we addressed the Lys_{33} $^{13}C_{33}^{\beta}$–$^{13}C_{33}^{\alpha}$–$^{13}C'_{33}$ three-spin system and the $^{13}C_i^{\alpha}$–$^{13}C_i^{\beta}$–$^{13}C_i^{\gamma}$–$^{13}C'_i$ four-spin systems of Asp_{32} and Glu_{34} to optimize the C7 transfers for the N(CO)CA (Fig. 4f), N(CA)CB (Fig. 4g), and N(CA)CX (Fig. 4h) correlations. The $^{13}C'$ to $^{13}C^{\alpha}$ transfer for the N(CO)CA experiment is readily optimized for selective transfer into the desired $^{13}C^{\alpha}$ spin (49% efficiency; $\tau_{C7} = 6\tau_r$ and $\omega_{\text{off}}^{C7}/2\pi = 1.5$ kHz) without noticeable leak further out in the side-chain because of the need for C7 to be close to the mean isotropic chemical shift of the two involved spins (*64, 67*). This condition cannot be fulfilled for the one-bond $^{13}C'$–$^{13}C^{\alpha}$ and $^{13}C^{\alpha}$–$^{13}C^{\beta}$ transfers simultaneously implying that the major source to $^{13}C^{\beta}$ coherence would be through the weak (long-range) $^{13}C'$–$^{13}C^{\beta}$ dipole–dipole interaction. In contrast, a similar kind of chemical shift truncation is more difficult to achieve for the ^{13}C–^{13}C transfer in the N(CA)CX-type of experiments (X aliphatic). In this case the homonuclear transfer will involve only aliphatic carbons, implying that all one-bond and long-range couplings will contribute to transfers between the carbons up to any terminating carbonyl (or similar) end group. This feature becomes clearly evident from Figs. 4g and 4h showing coherence transfers originating from the $^{13}C^{\alpha}$ spin in Asp_{32} and Glu_{34} and using the same C7 pulse sequence (optimally $\omega_{\text{RF}}^{C}/2\pi = 35$ kHz, $\tau_{C7} = 1.2$ ms, and $\omega_{\text{off}}^{C7}/2\pi = -8.5$ kHz) element for coherence transfer. In the former case the transfer is truncated by the C^{γ} carboxyl of Asp_{32} leading to transfer of -67% of the coherence to the desired $^{13}C^{\beta}$ spin in a N(CA)CB-type experiment, while 25% remains on the $^{13}C^{\alpha}$ spin (Fig. 4g). We note that the residual $^{13}C^{\alpha}$ coherence typically will not cause problems in the assignments due to the opposite sign and the significantly different isotropic chemical shift regimes for $^{13}C^{\alpha}$ and $^{13}C^{\beta}$. The same type of chemical shift truncation behavior is expected for Ala, Ser, Cys, and Asn and, depending on the selectivity of the pulse sequence, for Phe, Tyr, His, and Trp residues. In contrast, the same C7 building block will lead to N(CA)CX type of correlations for Val, Pro, Met, Ile, Leu, Glu, Lys, Arg, and Gln residues as demonstrated for the $^{13}C^{\delta}OO^{-}$-terminated Glu_{34} residue in Fig. 4h. In this case, the original $^{13}C^{\alpha}$ coherence is distributed with efficiencies of 25, -33, and 27% on the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{13}C^{\gamma}$ spins leading to an N(CA)CX type of correlation spectrum which through $^{13}C^{\beta}$ versus $^{13}C^{\alpha}$, $^{13}C^{\gamma}$ sign discrimination may prove very useful for residue identification. It is noted that the coherence will be distributed further out in the side chain for Ile, Leu, Lys, Arg, and Pro residues having aliphatic $^{13}C^{\delta}$ carbons.

Considering that the two pulse sequence elements will form consecutive elements in a 2D pulse sequence (Fig. 4a) operating on uniformly labeled peptides, it appears relevant to refine the optimization using a larger spin system and focusing on the overall efficiency of transfer to the desired destination spin(s). In this manner it will be possible to detect potential interplay between the pulse-sequence elements and multiple-spin effects. Thus, we conducted a 4D grid search over the ^{13}C carrier frequencies for the DCP and C7 pulse sequence elements and their durations (restricted to integrals of the rotor period) using the $^{13}C_{33}^{\beta}$–$^{13}C_{33}^{\alpha}$–$^{13}C'_{33}$–$^{15}N_{33}$–$^{13}C_{34}^{\alpha}$, $^{15}N_{32}$–$^{13}C_{32}^{\alpha}$–$^{13}C_{32}^{\beta}$–$^{13}C_{32}^{\gamma}$–$^{13}C'_{32}$, and $^{15}N_{34}$–$^{14}C_{34}^{\alpha}$–$^{13}C_{34}^{\beta}$–$^{13}C_{34}^{\gamma}$–$^{13}C'_{34}$ five-spin systems for the N(CO)CA, N(CA)CB, and N(CA)CX experiments, respectively, and the same RF field strength conditions as above. This optimization led to the same pulse sequences as obtained for the smaller spin systems above in terms of carrier offset frequencies and mixing times but slightly reduced overall efficiencies. For the N(CO)CA and N(CA)CB correlations the transfer efficiencies for the desired spin are 25.8 and 38.6%, respectively, to be compared with the values of 29.1 and 47.9% obtained by multiplication of the DCP and C7 transfer

efficiencies obtained above. For the Lys_{34} N(CA)CX correlation we obtained $\{^{13}C^{\alpha},^{13}C^{\beta},^{13}C^{\gamma}\}$ efficiencies of $\{-14.8, 23.3, -11.4\}$% compared with the multiplicated values of $\{-18.8, 24.5, -18.2\}$%. In all cases we observe that the pulse sequences optimized on three- and four-spin systems remain optimum while larger, and more realistic, spin systems cause reduced transfer efficiencies as may be attributed to diffusion of coherences to undesired spins or nondetectable longitudinal or multiple-quantum coherences. To further demonstrate this aspect we calculated the effect of the optimum pulse sequences on various heteronuclear seven-spin systems (cf., Fig. 4c) originating from the Lys_{33}-Glu_{34} (N(CO)CA, N(CA)CX) and Gln_{31}-Asp_{32} (N(CA)CB) residues of ubiquitin and leading to the overall transfer efficiencies illustrated in Figs. 4i–4k. These results clearly reveal that it is possible to (i) obtain a very high degree of selectivity in the coherence transfer processes forming the basis for heteronuclear correlation experiments and (ii) about double the sensitivity of the experiment relative to commonly used pulse sequences based on broadband and non-γ-encoded transfer elements which in an optimistic view give an overall transfer efficiency of 10–12% assuming a transfer efficiency of 30–35% for each element under consideration of multiple-spin effects (*19*).

Obviously, the SIMMOL and SIMPSON combination may also be used for simulation of spectra obtained for specific polypeptide structures using advanced solid-state NMR experiments. To demonstrate this aspect and visualize the heteronuclear correlations that may be obtained using the 2D experiments described above, Fig. 5 shows simulated DCP-C7 based N(CO)CA and N(CA)CX 2D spectra obtained for five residues of ubiquitin as highlighted by the peptide planes in the associated SIMMOL molecular graphics representations. Using the isotropic chemical shift values of Wand *et al.* (*50*) and under appropriate consideration of the transfer efficiencies in Figs. 4i–4k, the 2D spectra were obtained by coadding 2D spectra from five five-spin simulations involving $^{13}C^{\beta}_{i-1}$–$^{13}C^{\alpha}_{i-1}$–$^{13}C'_{i-1}$–$^{15}N_i$–$^{13}C^{\alpha}_i$ and $^{15}N_i$–$^{13}C^{\alpha}_i$–$^{13}C^{\beta}_i$–$^{13}C^{\gamma}_i$–$^{13}C'_i$ for the N(CO)CA and N(CA)CX experiment, respectively. The pulse sequences used 256 t_1 increments, $\Delta t_1 = \tau_r$, $\Delta t_2 = \tau_r/6$ (512 points), and the States–Haberkorn–Ruben (*68*) approach to achieve phase sensitive spectra. The resulting 2D spectra, each requiring 90 h of CPU time on a cluster of five 450-MHz Pentium PCs using SIMPSON's distributed clustering feature, clearly demonstrate the predicted selectively and relative intensity of the various resonances and reinforce the attractive sign-encoding of the various resonances in the N(CA)CX experiment as a means to discriminate potentially overlapping $^{13}C^{\beta}$ and $^{13}C^{\gamma}$ resonances.

Simulation of 2D PISA Wheel Spectra

For membrane proteins detailed information about the structure and membrane associated conformation may alternatively be obtained by static-sample solid-state NMR experiments on isotope-labeled peptides incorporated into laminar phospholipid bilayers oriented with the membrane normal parallel to the external magnetic field (*55*). In particular, it has proven useful to determine orientation-dependent ^{15}N chemical shifts and ^{1}H–^{15}N dipolar couplings for ^{15}N-labeled peptides using the PISEMA experiment (*45*). For uniformly labeled samples, this experiment gives rise to so-called PISA wheels from which it, under ideal conditions, is possible to assign the ^{15}N resonance directly to the structure and independently of this determine the orientation of α-helical structures relative to the bilayer normal (*69, 70*). Furthermore, it has recently been shown that transmembrane β-strands cause similar types of systematic patterns in the PISEMA spectra (*71*). These attractive features render the PISEMA type of experiments very interesting for structure determination simply for the reason that time-consuming establishment of accurate assignments potentially may be avoided. Before this conclusion can be taken, however, it is very important to investigate how ideal the helical structure and the pulse sequence must be to prevent wrong interpretations. The combination of SIMMOL and SIMPSON forms an ideal vehicle for investigating such effects by exploiting SIMMOL to establish tensor information for all relevant ^{1}H–^{15}N containing spin clusters and SIMPSON for numerically exact calculations of the response from the PISEMA experiments under consideration of multiple-spin and finite RF pulse effects.

To illustrate the influence of the secondary structure and finite RF pulses on the appearance of the PISA wheels, Fig. 6 shows a series of molecular structures and associate PISA wheels obtained using SIMMOL in combination with SIMPSON for numerically exact calculations and SIMMOL alone for analytical calculation of PISEMA experiments assuming ideal RF pulse conditions (*70*). Figures 6b and 6c show SIMPSON simulations for an ideal α-helix ($\phi, \psi = -65^{\circ}, -40^{\circ}$) tilted 20° relative to the magnetic field direction as shown in Fig. 6a. A "worst case" numerical exact simulation represented by the contours in Fig. 6b employs relatively low RF field strengths of 40 kHz (32.7 kHz) for ^{15}N (^{1}H) during the $\omega_1/2\pi$ evolution period and a decoupling RF field strength of 80 kHz in the direct dimension. The (N_i, H_i, H^{α}_i, H_{i-1}) spin systems in this simulation have been chosen as to allow evaluation of the effect of inefficient homonuclear dipole coupling of the protons. Comparison with an ideal SIMMOL-calculated PISA wheel (*70*) as represented by a solid line in this figure reveals differences of up to 0.5 kHz in the $\omega_1/2\pi$ and 2.5 ppm in the $\omega_2/2\pi$ dimension. However, the overall wheel shape is preserved and the spectral artifacts, especially around the center of the $\omega_1/2\pi$ dimension arising from imperfect spin locking and off-resonance effects, diminish as the RF field strengths are increased. This is illustrated in Fig. 6c where the RF field strengths are doubled to 80 kHz (65.3 kHz) for ^{15}N (^{1}H). The full-structure SIMMOL simulations in Figs. 6e, 6f, and 6h required approximately 2–12 s of CPU time on a 450-MHz Pentium PC while the SIMPSON four-spin simulations (Figs. 6b, 6c, and 6i) of the two-dimensional

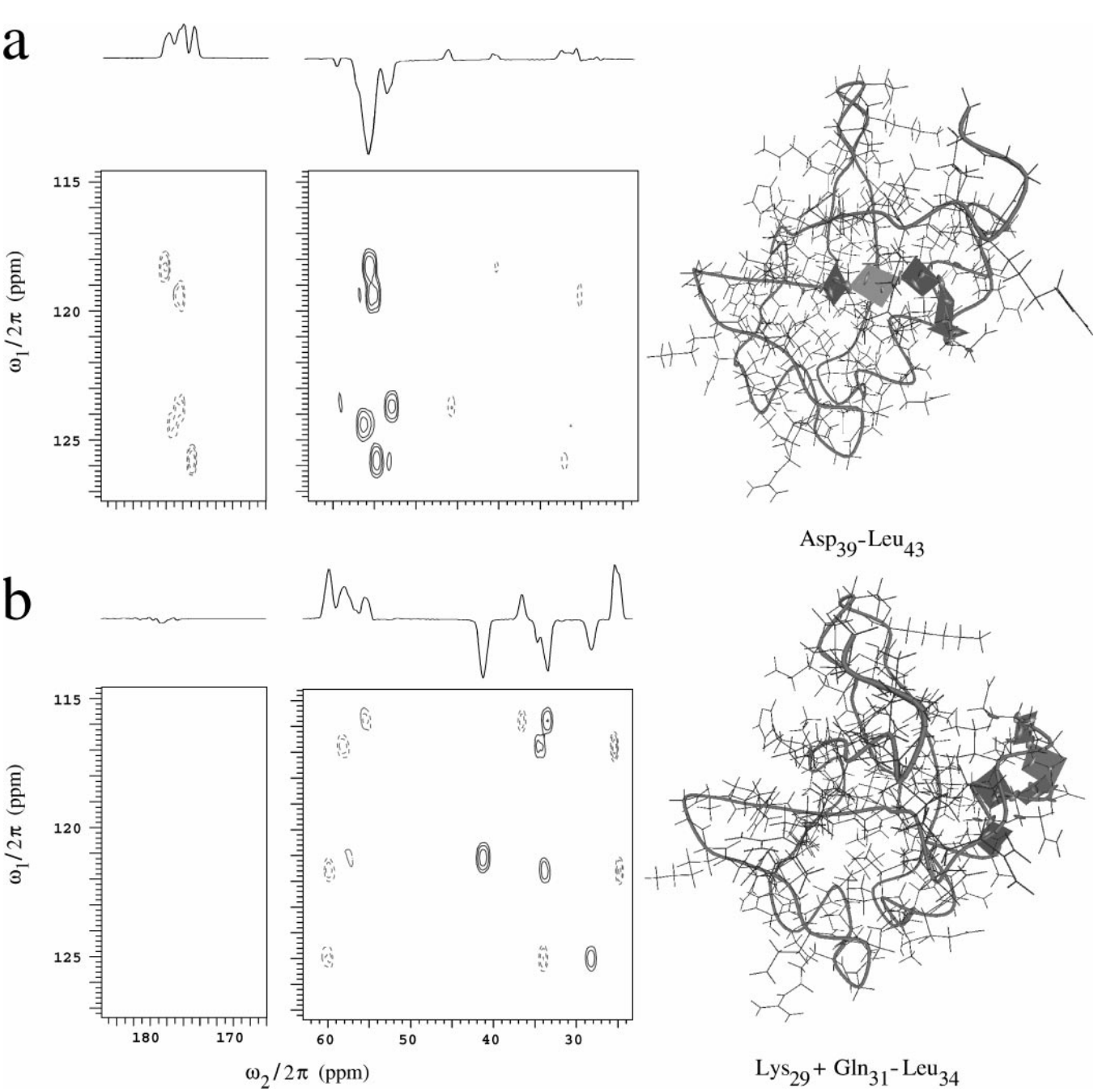

FIG. 5. Simulated DCP-C7 (a) N(CO)CA and (b) N(CA)CX 2D correlation spectra for selected residues in ubiquitin obtained using SIMPSON with tensor parameters for the involved spin systems established from the PDB file using SIMMOL along with the default chemical shift and J coupling parameters in Table 2. The simulations used $\omega_r/2\pi = 5$ kHz, $\omega_{RF}^{C7}/2\pi = \omega_{RF}^{DCP,C}/2\pi = 35$ kHz, $\omega_{RF}^{DCP,N}/2\pi = 30$ kHz, along with (a) $\tau_{DCP} = \tau_{C7} = 1.2$ ms and (b) $\tau_{DCP} = 1.8$ ms, $\tau_{C7} = 1.2$ ms. The ^{15}N carrier was on-resonance with respect to the amide nitrogens (120 ppm) while the ^{13}C carrier was set to the optimum values from Figs. 4i and 4k for N(CO)CA and N(CA)CX, respectively. The chemical shift, scalar J coupling, and dipolar coupling parameters used for the simulations can be found in the SIMPSON input files available at http://nmr.imsb.au.dk.

PISEMA spectra with 256×128 data points required 3–4 min per residue.

An impression of the known influence of secondary structure on the shape of the PISA wheel (*70*) is provided by the ideal wheel employing different α-helical torsion angles of $\phi, \psi = -57^\circ, -57^\circ$ as depicted in Fig. 6c by a dashed line. The apparent difference in the shape of this wheel relative to the wheel employing torsion angles of $\phi, \psi = -65^\circ, -40^\circ$ (solid line in Fig. 6c) indicates that PISA wheels are indeed very sensitive to changes in the secondary structure. To further investigate this aspect Figs. 6d–6i address attention to local variations in the torsion angles in the recent XRD bacteriorhodopsin (PDB ID:

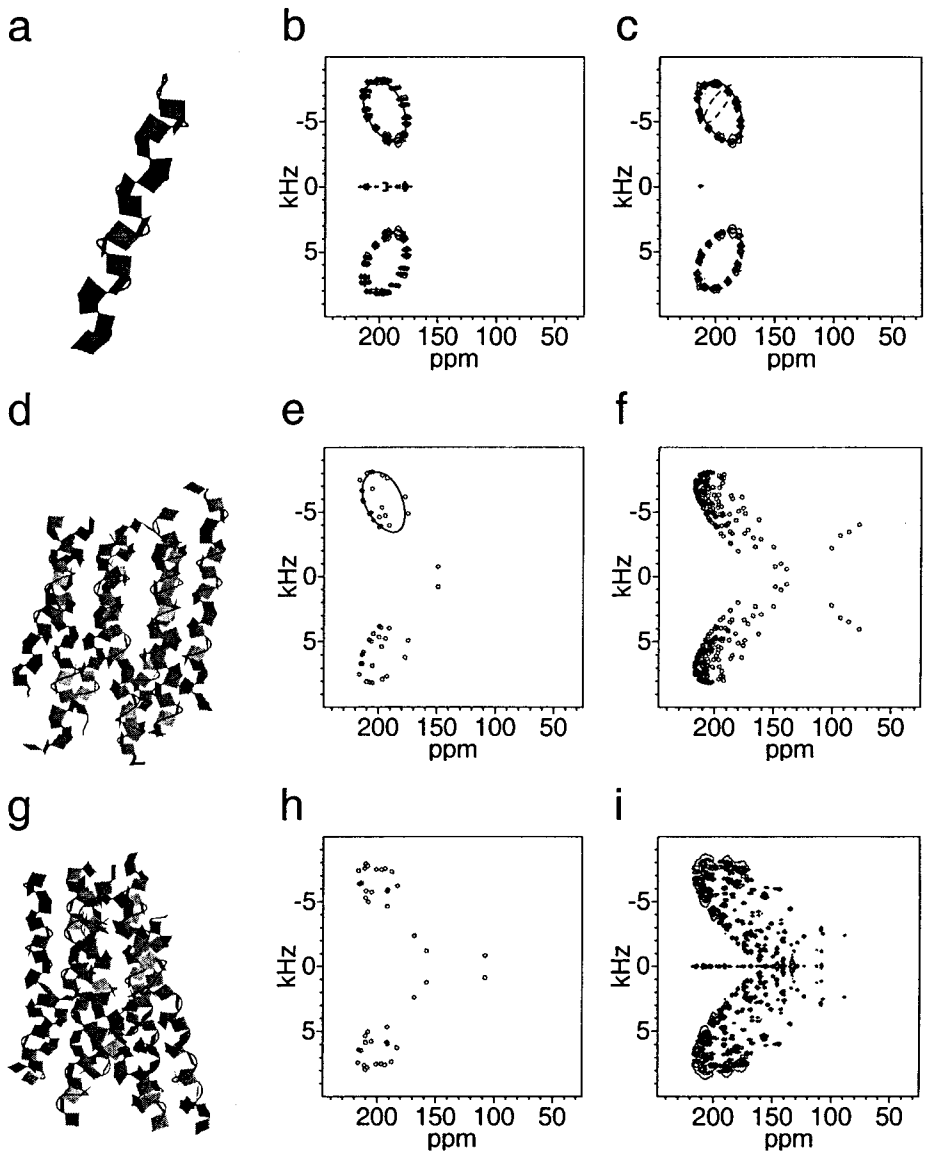

FIG. 6. Numerical simulations of PISEMA experiments (^{15}N shift horizontal; ^{1}H–^{15}N dipolar coupling vertical) on uniformly ^{15}N-labeled proteins using SIMMOL in combination with SIMPSON. (a) SIMMOL representation of an ideal 18-residue peptide with $\phi, \psi = -65^\circ, -40^\circ$. (b, c) PISEMA spectra for the peptide in (a) employing numerically exact simulations to account for finite RF pulse, off-resonance, and multiple-spin effects with ^{15}N SEMA RF field amplitudes ($\omega_{RF}^{N}/2\pi$) of (b) 40 and (c) 80 kHz (see text). The solid lines correspond to ideal PISA wheels. The dashed line in (c) corresponds to an ideal PISA wheel for an α helix with backbone torsion angles of $\phi, \psi = -57^\circ, -57^\circ$. (e, f) PISEMA spectra of helix 1 (e) and all seven transmembrane helices (f) of bacteriorhodopsin as visualized in (d) with helix 1 highlighted in black and marked by arrow. (h, i) Simulated PISEMA spectra of helix 4 (h) and all seven transmembrane helices (i) of rhodopsin as visualized in (g) with helix 4 highlighted in black and marked by an arrow. The latter spectrum (i) corresponds to a numerically exact simulation taking finite RF pulses, off-resonance, and multiple-spin effects into account (see text). The SIMMOL output and SIMPSON input files for the calculations are available at http://nmr.imsb.au.dk.

1C3W) (*72*) and rhodopsin (PDB ID: 1F88) (*73*) structures. The most ideal of the seven transmembrane (TM) helices of bacteriorhodopsin is helix 1 as judged from the standard deviation of the torsion angles. In Fig. 6d this helix is highlighted in black in the SIMMOL representation of all seven TM helices. The ideal PISEMA spectrum for helix 1 is shown in Fig. 6e and still reveals a wheel shape that fits fairly well with the wheel for an ideal helix tilted 20° relative to the magnetic field direction (represented by the solid line). However, an ideal PISEMA spectrum for all seven TM helices of bacteriorhodopsin (shown in Fig. 6f) hardly reflects any wheel shapes at all.

The difficulty for the topology-based interpretation becomes even more apparent by inspecting the PISEMA spectra calculated for the most perfect helix (Fig. 6h) and all seven TM helices (Fig. 6i) of rhodopsin visualized by SIMMOL in Fig. 6g. The PISEMA spectrum in Fig. 6i is calculated under

consideration of effects from finite RF pulses, off-resonance effects, and imperfect homonuclear decoupling of the protons. That is, the spectrum from each N–H spin pair results from a simulation of a three- to four-spin system using the same spin system as for the ideal helix in Fig. 6b (except for the absence of amide protons for PRO residues) but with ^{15}N (^{1}H) RF amplitudes of 50 kHz (40.8 kHz) for the SEMA block. The deviation from the ideal PISA wheel shape in the spectrum for the single helical fragment (Fig. 6h) indicates that it may be difficult to obtain precise structure information without assistance from complementary methods even for a single helical fragment of a membrane protein. Obviously, this situation becomes even more evident when several differently oriented helices contribute to the spectra as revealed by the "powder-like" PISEMA spectrum for all seven TM helices of rhodopsin in Fig. 6i. In this discussion, however, it is important to realize that there might be two origins to the distortions of the simulated PISA wheels, as pointed out by Cross *et al.* (*74*). One is local variations in the torsion angles of the true molecule structure; another is uncertainties in the measured molecular structure (i.e., in the PDB file) from which the present PISA wheels are calculated. Obviously, the latter problem should be particularly pronounced for structures refined to low resolution. Thus, this effect may contribute to the difference in helix ideality for bacteriorhodopsin and rhodopsin which are refined to 1.55 and 2.8 Å resolution, respectively.

CONCLUSION

In conclusion, we have introduced the SIMMOL program for fast and efficient establishment of anisotropic nuclear spin interaction parameters for numerical simulation of solid-state NMR experiments on polypeptides. In addition, the software enables 3D visualization of anisotropic tensors, internuclear axes, peptide planes, etc., in the molecular structure being useful in the setup of relevant spin systems and for presentation of solid-state (or partially aligned liquid-state) NMR results relying on anisotropic interactions. The software is designed to export information about the structure-dependent nuclear spin interactions directly into a format readable by the SIMPSON software package, a powerful and widely used tool for numerical simulations. With these features, it is foreseen that SIMMOL will form an indispensable supplementary tool to the SIMPSON software package for evaluation and application of solid-state NMR experiments for assignment and structure determination of polypeptides in the solid phase. Through its independent nature, SIMMOL can obviously also be used for visualization purposes alone or for the establishment of nuclear spin tensor information for use with other software than SIMPSON.

ACKNOWLEDGMENTS

This research was supported by grants from The Danish Research Agency in relation to the Danish Biotechnology Instrument Centre (J. No. 0002214), Carlsbergfondet (J. No. 990258/20-863), the Danish Natural Science Council (J. No. 9901954), Novo Nordisk Fonden, and the European Commission (BIO4-CT97-2101).

REFERENCES

1. U. Haeberlen and J. S. Waugh, Coherent averaging effects in magnetic resonance, *Phys. Rev.* **175,** 453–467 (1968).
2. M. Hohwy and N. C. Nielsen, Systematic design and evaluation of multiple-pulse experiments in nuclear magnetic resonance spectroscopy using a semi-continuous Baker–Campbell–Hausdorff expansion, *J. Chem. Phys.* **109,** 3780–3791 (1998).
3. H. Liu, S. J. Glaser, and G. Drobny, Development and optimization of multiple pulse propagators. Applications to homonuclear spin decoupling in solids, *J. Chem. Phys.* **93,** 7543–7560 (1990).
4. D. Sakellariou, A. Lesage, P. Hodgkinson, and L. Emsley, Homonuclear dipolar decoupling in solid-state NMR using continuous phase modulation, *Chem. Phys. Lett.* **319,** 253–260 (2000).
5. T. Gullion and J. Schaefer, Rotational-echo double resonance NMR, *J. Magn. Reson.* **81,** 196–200 (1989).
6. J. Skibsted, N. C. Nielsen, H. Bildsøe, and H. J. Jakobsen, Satellite transitions in MAS NMR spectra of quadrupolar nuclei, *J. Magn. Reson.* **95,** 88–117 (1991).
7. F. H. Larsen, H. J. Jakobsen, P. D. Ellis, and N. C. Nielsen, Sensitivity-enhanced quadrupolar-echo NMR of half-integer quadrupolar nuclei. Magnitudes and relative orientation of chemical shielding and quadrupolar coupling tensors, *J. Phys. Chem. A* **101,** 8597–8606 (1997).
8. S. Dusold, E. Klaus, A. Sebald, M. Bak, and N. C. Nielsen, Magnitudes and relative orientations of chemical shielding, dipolar and *J* coupling tensors for isolated ^{31}P–^{31}P spin pairs determined by iterative fitting of ^{31}P MAS NMR spectra, *J. Am. Chem. Soc.* **119,** 7121–7129 (1997).
9. X. Feng, Y. K. Lee, D. Sandström, M. Edén, H. Maisel, A. Sebald, and M. H. Levitt, Direct determination of molecular torsion angle by solid-state NMR, *Chem. Phys. Lett.* **257,** 314–320 (1996).
10. D. P. Weliky and R. Tycko, Determination of peptide conformations by two-dimensional magic-angle spinning NMR exchange spectroscopy with rotor synchronization, *J. Am. Chem. Soc.* **118,** 8487–8488 (1996).
11. M. Hong, J. D. Gross, and R. G. Griffin, Site-resolved determination of peptide torsion angle ϕ from relative orientations of backbone N–H and C–H bonds by solid-state NMR, *J. Phys. Chem. A* **101,** 5869–5874 (1997).
12. T. Fujiwara, T. Shimomura, Y. Ohigashi, and H. Akutsu, Multidimensional solid-state nuclear magnetic resonance for determining the dihedral angle from correlation of ^{13}C–^{1}H and ^{13}C–^{13}C dipolar interactions under magic-angle spinning conditions, *J. Chem. Phys.* **109,** 2380–2393 (1998).
13. M. Hohwy, C. P. Jaroniec, B. Reif, C. M. Rienstra, and R. G. Griffin, Local structure and relaxation in solid-state NMR: Accurate measurement of amide N–H bond lengths and H–N–H bond angles, *J. Am. Chem. Soc.* **122,** 3218–3219 (2000).
14. M. Edén, A. Brinkmann, H. Luthman, L. Erikson, and M. H. Levitt, Determination of molecular geometry by high-order multiple-quantum evolution in solid-state NMR, *J. Magn. Reson.* **144,** 266–279 (2000).
15. L. Odgaard, M. Bak, H. J. Jakobsen, and N. C. Nielsen, ^{13}C chemical shift and ^{13}C–^{14}N dipolar coupling tensors determined by ^{13}C rotary resonance solid-state NMR, *J. Magn. Reson.* **148,** 298–308 (2001).
16. M. Kainosho, Isotope labeling of macromolecules for structural determinations, *Nat. Struct. Biol.* **4,** 858–861 (1997).
17. M. Bak, J. T. Rasmussen, and N. C. Nielsen, SIMPSON: A general simulation program for solid-state NMR spectroscopy, *J. Magn. Reson.* 147, 296–330 (2000). Internet address and download site: http://nmr.imsb.au.dk.

18. J. W. Eaton and others, GNU Octave, A high-level interactive language for numerical computations, 3rd Ed., 1997. Open source software at http://www.octave.org.
19. M. Bak, R. Schultz, and N. C. Nielsen, Numerical simulations for experiment design and extraction of structural parameters in biological solid-state NMR spectroscopy, *in* "Perspectives on Solid-State NMR in Biology" (S. Kiihne and H. J. M. de Groot, Eds.), pp. 95–109, Kluwer Academic, Dordrecht, 2001.
20. B. B. Welch, "Practical Programming in Tcl and Tk," Prentice Hall, Englewood cliffs, NJ 1995. The open source Tcl/Tk software can be downloaded via, e.g., the Tcl Developers Xhange homepage http://dev.scriptics.com.
21. R. A. Engh and R. Huber, Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr.* **A47,** 392–400 (1991).
22. G. E. Schulz and R. H. Schirmer, "Principles of Protein Structure," Springer-Verlag, New York, 1988.
23. The GNU General Public License is available from the Free Software Foundation, Boston, MA, at http://www.gnu.org.
24. S. Levy, T. Munzner, M. Philips, and others, Geomview 1.6.1, University of Minnesota, Minneapolis. 1996. Open source software at http://www.geomview.org.
25. F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. E. Meier, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, The Protein Data Bank: a computer-based archival file for macromolecular structures, *J. Mol. Biol.* **112,** 535–542 (1977). Internet address: http://www.rcsb.org/pdb.
26. G. Vriend, WHAT IF: A molecular modelling and drug design program. *J. Mol. Graph.* **8,** 52–56 (1990). Internet address: http://www.cmbi.kun.nl/whatif.
27. J. Kyte and R. F. Doolittle, A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132 (1982).
28. S. J. Opella, P. L. Stewart, and K. G. Valentine, Protein structure by solid-state NMR spectroscopy. *Q. Rev. Biophys.* **19,** 7–49 (1987).
29. D. J. Siminovitch, Solid-state NMR studies of proteins: The view from static ^{2}H NMR experiments. *Biochem. Cell Biol.* **76,** 411–422 (1998).
30. T. A. Cross and J. R. Quine, Protein structure in anisotropic environments: Development of orientational constraints. *Concepts Magn. Reson.* **12,** 55–70 (2000).
31. C. H. Wu, A. Ramamoorthy, L. M. Gierasch, and S. J. Opella, Simultaneous characterization of the amide ^{1}H chemical shift and ^{1}H–^{15}N dipolar and ^{15}N chemical shift interaction tensors for a peptide bond by three-dimensional solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **117,** 6148–6149 (1995).
32. A. Naito, S. Ganapathy, K. Akasaka, and C. A. McDowell, Chemical shielding tensor and ^{13}C–^{14}N dipolar splitting in single crystals of L-alanine, *J. Chem. Phys.* **74,** 3190–3197 (1981).
33. Y. Wei, D. Lee, and A. Ramamoorthy, Solid-State ^{13}C NMR chemical shift anisotropy tensors of polypeptides, *J. Am. Chem. Soc.* **123,** 6118–6126 (2001).
34. C. J. Hartzell, M. Whitfield, T. G. Oas, and G. P. Drobny, Determination of the ^{15}N and ^{13}C chemical shift tensors of L-[^{13}C]–L-[^{15}N]alanine from the dipole-coupled powder patterns, *J. Am. Chem. Soc.* **109,** 5966–5969 (1987).
35. Q. Teng, M. Iqbal, and T. A. Cross, Determination of the ^{13}C chemical shift and ^{14}N electric field gradient tensor orientations with respect to the molecular frame in a polypeptide. *J. Am. Chem. Soc.* **114,** 5312–5321 (1992).
36. W. M. Tan, Z. Gu, A. C. Zeri, and S. J. Opella, Solid-state NMR triple-resonance backbone assignments in a protein, *J. Biomol. NMR* **13,** 337–342 (1999).
37. D. Lee, Y. Wei, and A. Ramamoorthy, A two-dimensional magic-angle decoupling and magic-angle turning solid-state NMR method: An application to study chemical shift tensors from peptides that are nonselectively labeled with ^{15}N isotope, *J. Phys. Chem. B* **105,** 4752–4762 (2001).
38. S. Schramm and E. Oldfield, Nuclear magnetic resonance studies of amino-acids and proteins. Rotational correlation times of proteins by deuterium nuclear magnetic resonance spectroscopy, *Biochemistry* **22,** 2908–2913 (1983).
39. A. W. Hing, S. P. Adams, D. F. Silbert, and R. E. Norberg, Deuterium NMR of Val^{1} $\cdots$ (2-^{2}H)Ala^{3} $\cdots$ gramicidin A in oriented DMPC bilayers, *Biochemistry* **29,** 4144–4156 (1990).
40. R. S. Prosser, J. H. Davies, F. W. Dahlquist, and M. A. Lindorfer, ^{2}H nuclear magnetic resonance of the gramicidin A backbone in a phospholipid bilayer. *Biochemistry* **30,** 4687–4696 (1991).
41. R. E. Stark, R. A. Haberkorn, and R. G. Griffin, ^{14}N NMR determination of NH bond lengths in solids, *J. Chem. Phys.* **68,** 1996–1997 (1978).
42. K. Yamauchi, S. Kuroki, I. Ando, T. Ozaki, and A. Shoji, ^{17}O chemical shifts and quadrupole coupling constants in poly(L-alanine)s determined using a high-speed MAS technique. *Chem. Phys. Lett.* **302,** 331–336 (1999).
43. R. H. Havlin, H. Le, D. D. Laws, A. C. de Dios, and E. Oldfield, An ab initio quantum chemical investigation of carbon-13 NMR shielding tensors in glycine, alanine, valine, isoleucine, serine, and threonine: Comparisons between helical and sheet tensors, and the effect of χ_1 on shielding *J. Am. Chem. Soc.* **119,** 11951–11958 (1997).
44. M. Sattler, J. Schleucher, and C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients, *Progr. NMR Spectrosc.* **34,** 93–158 (1999).
45. C. H. Wu, A. Ramamoorthy, and S. J. Opella, High-resolution heteronuclear dipolar solid-state NMR spectroscopy, *J. Magn. Reson. A* **109,** 270–272 (1994).
46. H. W. Spiess, Rotation of molecules and nuclear spin relaxation, *NMR Basic Prin. Progr.* **15,** 55–214 (1978).
47. M. Mehring, "High-Resolution NMR in Solids," Springer-Verlag, New York, 1983.
48. Y. Wei, A. C. de Dios, and A. E. McDermott, Solid-state ^{15}N NMR chemical shift anisotropy of histidines: Experimental and theoretical studies of hydrogen bonding, *J. Am. Chem. Soc.* **121,** 10389–10394 (1999).
49. G. Cornilescu, J. S. Marquardt, M. Ottiger, and A. Bax, Validation of protein structure from anisotropic carbonyl chemical shifts in dilute liquid crystalline phase, *J. Am. Chem. Soc.* **120,** 6836–6837 (1998).
50. A. J. Wand, J. L. Urbauer, R. P. McEvoy, and R. J. Bieber, Internal dynamics of human ubiquitin revealed by ^{13}C-relaxation studies of randomly fractionally labeled protein, *Biochemistry* **35,** 6116–6125 (1996).
51. J. Gesell, M. Zasloff, and S. J. Opella, Two-dimensional ^{1}H NMR experiments show that the 23-residue magainin antibiotic peptide is an α-helix in dodecylphosphocholine micelles, sodium dodecylsulfate micelles, and trifluoroethanol/water solution. *J. Biomol. NMR* **9,** 127–135 (1997).
52. J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegaard, Nuclear magnetic dipole interactions in field-oriented proteins: Information for structure determination in solution, *Proc. Natl. Acad. Sci. USA* **92,** 9279–9283 (1995).
53. M. Tjandra and A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, *Science* **278,** 1111–1114 (1997).
54. R. G. Griffin, Dipolar recoupling in MAS spectra of biological solids, *Nat. Struct. Biol.* **5,** 508–512 (1998).
55. S. J. Opella, NMR of membrane proteins, *Nat. Struct. Biol.* **4,** 845–848 (1997).
56. R. Tycko, Prospects for resonance assignments in multidimensional solid-state NMR spectra of uniformly labeled proteins, *J. Biomol. NMR* **8,** 239–251 (1996).

57. S. K. Straus, T. Bremi, and R. R. Ernst, Experiments and strategies for the assignment of fully $^{13}C/^{15}N$-labelled polypeptides by solid-state NMR, *J. Biomol. NMR* **12,** 39–50 (1998).

58. M. Hong and R. G. Griffin, Resonance assignments for solid peptides by dipolar-mediated $^{13}C/^{15}N$ correlation solid-state NMR, *J. Am. Chem. Soc.* **120,** 7113–7114 (1998).

59. M. Hong, Resonance assignment of $^{13}C/^{15}N$ labeled solid proteins by two- and three-dimensional magic-angle-spinning NMR, *J. Biomol. NMR* **15,** 1–14 (1999).

60. A. McDermott, T. Polenova, A. Bockmann, K. W. Zilm, E. K. Paulsen, R. W. Martin, and G. T. Montelione, Partial NMR assignments for uniformly (^{13}C, ^{15}N)-enriched BPTI in the solid state, *J. Biomol. NMR* **16,** 209–219 (2000).

61. J. Pauli, M. Baldus, H. J. M. de Groot, and H. Oschinat, Backbone and side chain ^{13}C and ^{15}N signal assignments of the α-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 Tesla, *Chembiochem* **2,** 272–281 (2001).

62. C. M. Rienstra, M. Hohwy, M. Hong, and R. G. Griffin, 2D and 3D ^{15}N–^{13}C–^{13}C NMR chemical shift correlation spectroscopy of solids: Assignment of MAS spectra of peptides, *J. Am. Chem. Soc.* **122,** 10979–10990 (2000).

63. J. Schaefer, E. O. Stejskal, J. R. Garbow, and R. A. McKay, Quantitative determination of the concentrations of ^{13}C–^{15}N chemical bonds by double cross-polarization NMR, *J. Magn. Reson.* **59,** 150–156 (1984).

64. Y. K. Lee, N. D. Kurur, M. Helmle, O. G. Johannessen, N. C. Nielsen, and M. H. Levitt, Efficient dipolar recoupling in the NMR of rotating solids. A sevenfold symmetric radiofrequency pulse sequence, *Chem. Phys. Lett.* **242,** 304–309 (1995).

65. N. C. Nielsen, H. Bildsøe, H. J. Jakobsen, and M. H. Levitt, Double-quantum homonuclear rotary resonance: Efficient dipolar recovery in magic-angle spinning nuclear magnetic resonance, *J. Chem. Phys.* **101,** 1805–1812 (1994).

66. C. M. Rienstra, M. E. Hatcher, L. J. Mueller, B.-Q. Sun, S. W. Fesik, and R. G. Griffin, Efficient multispin homonuclear double-quantum recoupling for magic-angle spinning NMR: ^{13}C–^{13}C correlation spectroscopy of U–^{13}C-erythromycin A, *J. Am. Chem. Soc.* **120,** 10602–10612 (1998).

67. M. Hohwy, H. J. Jakobsen, M. Edén, M. H. Levitt, and N. C. Nielsen, Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence, *J. Chem. Phys.* **108,** 2686–2694 (1998).

68. D. J. States, R. A. Haberkorn, and D. J. Ruben, A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants, *J. Magn. Reson.* **48,** 286–292 (1982).

69. F. Marassi and S. J. Opella, A solid-state NMR index of helical membrane protein structure and topology, *J. Magn. Reson.* **144,** 150–155 (2000).

70. J. Wang, J. Denny, C. Tian, S. Kim, Y. Mo, F. Kovacs, Z. Song, K. Nishimura, Z. Gan, R. Fu, J. R. Quine, and T. A. Cross, Imaging membrane protein helical wheels, *J. Magn. Reson.* **144,** 162–167 (2000).

71. F. M. Marassi, A simple approach to membrane protein secondary structure and topology based on NMR spectroscopy, *Biophys. J.* **80,** 994–1003 (2001).

72. H. Luecke, B. Schobert, H.-T. Richter, J.-P. Cartailler, and J. K. Lanyi, Structure of bacteriorhodopsin at 1.55 Angstrom resolution, *J. Mol. Biol.* **291,** 899–911 (1999).

73. K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, L. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, and M. Miyano, Crystal structure of rhodopsin: A G protein-coupled receptor, *Science* **289,** 739–745 (2000).

74. T. A. Cross, S. Kim, J. Wang, R. Fu, and J. Quine, From topology to high resolution membrane protein structures, *in* "Perspectives on Solid-State NMR in Biology" (S. Kiihne and H. J. M. de Groot, Eds.), pp. 55–69, Kluwer Academic, Dordrecht, 2001.