Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 186 (2007) 34–50
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

A versatile component-coupling model to account for substituent effects: Application to polypeptide ϕ and χ_1 torsion related ^3J data

Jürgen M. Schmidt *

Biosciences Department, University of Kent, Canterbury CT2 7NJ, UK

Received 17 July 2006; revised 14 January 2007
Available online 23 January 2007

Abstract

A model is proposed for collating fundamental and incremental component couplings to account for substituent effects on ^3J arising from, for example, amino-acid type variation. The unique topology patterns encountered in each of the common amino acids were modeled by assigning substituents on a ^3J coupling path to four simple categories comprising only relative positions: central (*inner*) *vs*. terminal (*outer*) and first-sphere *vs*. second-sphere. Associated increment values then reflect the influences on each ^3J coupling accessible for torsion-angle determination. Facility of use of this model, in comparison with previous ones, owes to its strict limitation to no more than three Karplus coefficients for each specific torsion-angle dependency derived. The model was integrated in the concept of self-consistent ^3J analysis and applied to polypeptide fragments X–N–C^α–Y and X–C^α–C^β–Y related to torsions ϕ and χ_1, respectively, yielding quantitative effects of both first- and second-sphere substituents. Regarding the polypeptide backbone, the model predicts first-sphere substituent effects on ϕ-related ^3J couplings to be within experimental uncertainty because main-chain topologies are identical in most amino-acid types, except for marginal effects in glycine and proline. However, effects in excess of standard errors in $^3J(\phi)$ measurements are anticipated from second-sphere substituent variation. Regarding amino-acid side chains, first-sphere substituent effects on χ_1-related ^3J couplings were previously found pivotal to accurate torsion-angle interpretation. Taking additional second-sphere effects on $^3J(\chi_1)$ into account is here demonstrated further to improve biomolecular structure analysis.

Keywords: Vicinal coupling constants; Torsion angle; Data redundance; Self-consistency; Flavodoxin

1. Introduction

Torsion angles derived from NMR ^3J coupling constants help restrict the accessible conformational space when building three-dimensional molecular models [1]. Rules for translating between ^3J and dihedral angles, θ, rest on so-called Karplus coefficients [2], in essence, multipliers to the cosine terms obtained when developing the coupling constant as a Fourier series in the angle. Derived from quantum-mechanical calculations of the dihedral-angle dependence of overlap integrals [3], the Karplus relation presents a concise form, solely involving three variable parameters per torsion-coupling pair, supplemented by a fixed phase increment to the torsion, $\Delta\theta$, that depends on the angle type in question. Traditionally, Karplus coefficients for studies on biomolecules are being derived with reference to high-resolution X-ray structures [4].

^3J analyses are often carried out qualitatively. Given a ^3J coupling constant, the Karplus relation confines solutions to the sought torsion angle to up to four θ values. Unique and unambiguous values can only be obtained from multiple coupling constants collected for the same torsion, commanding the corresponding sets of Karplus coefficients be available. However, Karplus coefficients are not always available or sufficiently accurate for the particular combination of *J*-coupling type and molecular fragment in question, discouraging routine measurement of redundant ^3J datasets.

Pioneering work accomplished quantification of the influence of substituents on ^{1}H,^{1}H-homonuclear ^3J cou-

* Fax: +44 1227 763912.
E-mail address: j.m.schmidt@kent.ac.uk

1090-7807/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2007.01.012

plings in five-membered proline and ribose ring systems [5,6]. To account for contributions from direct and remote substituents, sets of up to seven coefficients were established. It is emphasized here that respective coupling paths, ^{1}H–C–C–^{1}H or ^{1}H–N–C–^{1}H, allow the spin topology to vary at the two intervening positions only.

Affordable isotopic ^{15}N and ^{13}C labeling of biomolecular samples now allows one to access every conceivable ^3J coupling in proteins. Departure from probing ^{1}H spins only paves the way for a plethora of potential substituents bonded no longer only to the two intervening sites in the respective four-atom fragment, but also to the terminal, actively coupled nuclei themselves. The sheer number of possible substituent patterns accompanying heteronuclear couplings defied systematic quantification of effects because of difficulty in finding suitable compounds with known reference geometry.

Provided an excess of ^3J coupling data, redundant ^3J information was demonstrated to give accurate torsion angles without any prior knowledge of Karplus coefficients [7,8]. In the concept, both torsion angles and Karplus coefficients were self-consistently adjusted in an iterative procedure, obviating the otherwise critical need for highly resolved reference coordinates. A total of more than 1500 homo- and heteronuclear ^3J coupling constants collected for the 147-residue protein *Desulfovibrio vulgaris* flavodoxin [7–10] enabled accurate determination of self-consistent torsion angles ϕ and χ_1 in the protein backbone [7] and side chains [8], respectively. The chance of genuine discrepancies between solution and crystal states notwithstanding, likely as a result of different experimental conditions, convergence was achieved between torsion-angle values independently derived from both NMR spectroscopy and high-resolution X-ray crystallography [11–13], demonstrating the predictive power of the approach.

Redundant structure information self-consistently exploited allows one to probe influences on ^3J couplings other than the principal dihedral-angle dependencies. Among the pertinent effects are angular dynamics [14–18] as well as variations in the local spin topology [4,19–21]. The quantitative flavodoxin study [8] corroborated common qualitative experience that side-chain related coupling constants are affected indeed by topological variation across different amino-acid types, such as to cause misinterpretation of torsion angles if disregarded. So far, the initial application took into account effects from only those first-sphere substituents directly attached to either of the four atoms establishing the coupling path. Here, the self-consistent approach is extended and the incremental-coupling concept taken further, demonstrating that results improve when considering second-sphere substituents also.

2. Component-coupling concept

The empirical measure of substituent effects in a given coupling path is the coupling increment, $\Delta J^{\mathrm{R}}_{\mathrm{XY}} = J^{\mathrm{R}}_{\mathrm{XY}} - J^{\mathrm{H}}_{\mathrm{XY}}$, inferred from the experimental coupling constants $J^{\mathrm{R}}_{\mathrm{XY}}$ and $J^{\mathrm{H}}_{\mathrm{XY}}$ taken from substituted and unsubstituted compounds, respectively, denoting the change in the coupling J_{XY} on substituting a heavy atom for a notional hydrogen atom [19]. Incremental coupling constants, also known as component couplings, depend on both the type of a substituent and its position relative to the interacting coupled nuclei [20,21]. Two types of substituents shall be distinguished. The type '*inner*' encompasses those bonded to central atoms N′ and C$^\alpha$ in ϕ torsions, or to C$^\alpha$ and C$^\beta$ in χ_1 torsions, whereas '*outer*' encompasses those attached to terminal sites X and Y. Within these categories, relative positions are not being discriminated, so as to consider identical the effects from substituents bonded to either X or Y directed portions of the coupling path.

2.1. Relative scaling of substituent effects

Incremental Karplus coefficients are to reflect type and position of substituents rather than to depend on the type of the basic X–Y coupling. Consequently, substituent-specific increments to J shall apply uniformly to *all* coupling types connected with a particular torsion-angle type. Yet, coupling magnitudes do vary with J coupling type, such as to necessitate scaling of incremental effects. For example, $^3J_{\mathrm{H}\alpha,\mathrm{C}\gamma}$ coupling constants in amino acids usually exceed values of $^3J_{\mathrm{N}',\mathrm{C}\gamma}$ on an absolute measure, such that a substituent, oxygen attached to the outer C$^\gamma$ atom, say, is expected to impact on the former coupling more than on the latter. Similarly, inner C$^\beta$ bound substituents would affect the typically larger $^3J_{\mathrm{H}\alpha,\mathrm{H}\beta}$ couplings more than the comparatively smaller $^3J_{\mathrm{N}',\mathrm{H}\beta}$ couplings.

The angle-independent Karplus coefficient C_0 (coefficient A in Ref. [2]) identifies the mean J intensity obtained on complete torsion revolution and appears to correlate with the product gyromagnetic ratio of the coupled nuclei X and Y [22]. Previous studies [8] found substituents significantly to affect only this angle-invariant portion of the coupling interaction, essentially modulating the baseline value of J. Incremental effects, ΔC_0, any given substituent may have on the various X–Y pair combinations, are then being normalized using scaling factors defined here as

$$\gamma_{\mathrm{XY}} = (\gamma_{\mathrm{X}}\gamma_{\mathrm{Y}})^{1/2}(\gamma_{\mathrm{H}})^{-1}, \tag{1}$$

which yield 1.0000, 0.5014, 0.3183, 0.2515 and 0.1596 for interactions ^{1}H–^{1}H, ^{1}H–^{13}C, ^{1}H–^{15}N, ^{13}C–^{13}C and ^{13}C–^{15}N, respectively, disregarding the negative sign of γ_{N} for the purpose of numerical simplicity.

Our J-coupling model further uses increment frequencies, f_{R}, of any given substituent type R bound to the four-atom fragment, giving rise to a total coupling increment of

$$\Sigma\Delta J^{\mathrm{R}}_{\mathrm{XY}} = \gamma_{\mathrm{XY}}\Sigma f_{\mathrm{R}}\Delta C^{\mathrm{R}}_0, \tag{2}$$

where the sum runs over all relevant substituents. The extended Karplus equation is then given by [8]

$$^3J_{\mathrm{XY}}(\theta) = C_0 + C_1\cos\theta + C_2\cos 2\theta + \Sigma\Delta J^{R}_{\mathrm{XY}}. \tag{3}$$

2.2. First-sphere substituent patterns

Fig. 1 exemplifies the compilation of amino-acid dependent substituent counts for first-sphere substituent patterns related to the side-chain torsion χ_1 in asparagine. Eventually, all amino-acid topologies are distinguished solely through variation in their increment-count patterns. Arranging increment counts related to the six coupling types in blocks as seen in Table 1, the 3-by-2 layout of which also maps onto the panels in Fig. 1, visualises the construction of amino-acid topologies.

For example, switching from the series involving H^{β} as a coupling partner ($^3J_{X,H\beta}$ blocks on the left of Table 1) to those involving C^{γ} ($^3J_{X,C\gamma}$ blocks on the right) necessitates, in asparagine, inserting increments related to the terminal amide-group atoms, i.e., $O^{\delta 1}$ and $N^{\delta 2}$ bonded to C^{γ}. Also,

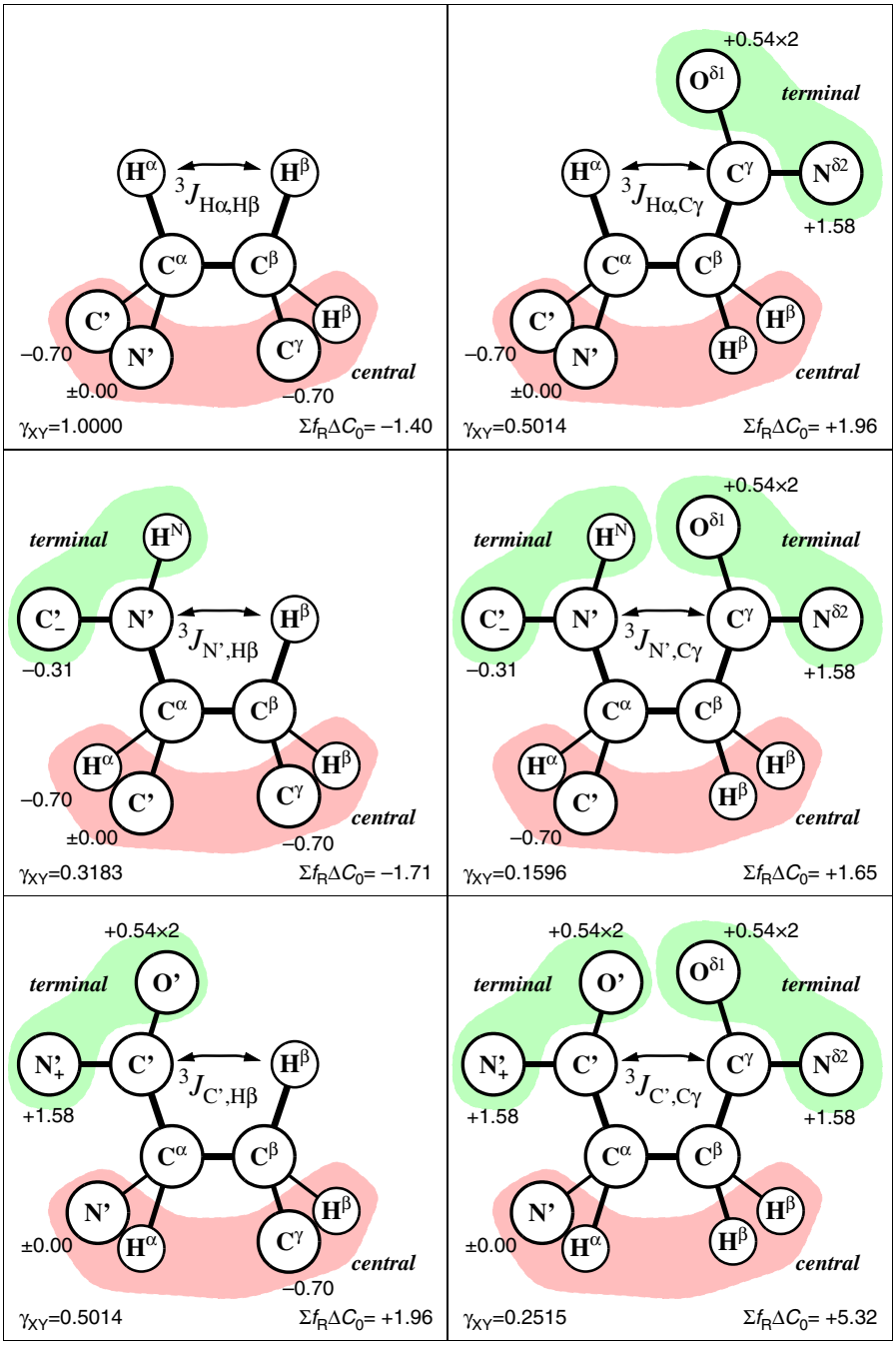

Fig. 1. Collation of first-sphere substituent effects on ^3J couplings as exemplified for the asparagine side-chain torsion χ_1 (see text for details). The scheme outlined is to be exercised separately for each of the common amino-acid types, resulting in the increment-count Table 1. Calculations using experimental data are then needed to determine the size of each coupling increment. Relative gyromagnetic ratios as per Eq. (1) are given on the lower left of each panel. These are to be multiplied with the cumulative component-coupling increments given on the lower right, as per Table 3, in order to estimate the total amino-acid specific change, ΔJ_0, in each respective coupling constant $^3J_{XY}$.

Table 1
Substituent frequencies for amino-acid specific ^3J coupling dependences on the torsion χ_1[a]

Type	Sphere	R	Thr	Ser	Cys	Ile	Val	Tyr	Trp	Phe	Leu	Met	Pro	Glu	Gln	Arg	Lys	His	Ala	Asp	Asn	Thr	Ser	Cys	Ile	Val	Tyr	Trp	Phe	Leu	Met	Pro	Glu	Gln	Arg	Lys	His	Ala	Asp	Asn
			$^3J_{H\alpha,H\beta}$																			$^3J_{H\alpha,C\gamma}$																			
Inner/central	1st	C	2	1	1	3	3	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	2nd	C	1	1	1	2	1	4	4	4	2	1	3	2	2	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1
		N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	5	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		S	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Outer/terminal	1st	C	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	3	3	3	2	—	1	1	1	1	1	2	—	—	—	—
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	1
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3	2	
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—
	2nd	C	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	2	2	—	1	—	—	—	—	1	1	—	—	—	—
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1	1	—	1	—	—	—	—
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	3	2	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
			$^3J_{N',H\beta}$																			$^3J_{N',C\gamma}$																			
Inner/central	1st	C	2	1	1	3	3	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		O	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	2nd	C	—	—	—	1	—	3	3	3	1	—	2	1	1	1	1	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	5	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		S	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Outer/terminal	1st	C	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	1	4	4	4	3	1	3	2	2	2	2	3	1	1	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	1
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3	2	
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—
	2nd	C	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	3	1	2	1	1	1	1	2	2	1	1	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1	1	—	1	—	—	—	—
		O	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	4	5	2	2	2	2	2	2
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

(*continued on next page*)

Table 1 (*continued*)

Type	Sphere	R	$^3J_{C',H\beta}$ Thr	Ser	Cys	Ile	Val	Tyr	Trp	Phe	Leu	Met	Pro	Glu	Gln	Arg	Lys	His	Ala	Asp	Asn	$^3J_{C',C\gamma}$ Thr	Ser	Cys	Ile	Val	Tyr	Trp	Phe	Leu	Met	Pro	Glu	Gln	Arg	Lys	His	Ala	Asp	Asn
Inner/central	1st	C	1	—	—	2	2	1	1	1	1	1	1	1	1	1	1	1	—	1	1	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	2nd	C	1	1	1	2	1	4	4	4	2	1	3	2	2	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3	2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—
Outer/terminal	1st	C	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	3	3	3	2	—	1	1	1	1	2	—	—	—	—
		N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	1	1	2
		O	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	5	4
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—
	2nd	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	3	1	2	1	1	1	2	2	1	1	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1	1	—	1	—	—	—
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	3	2	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

[a] *Notes:* Counts f_R for use with the extended Karplus equation $^3J(\theta) = C_0 + C_1\cos\theta + C_2\cos 2\theta + \gamma_{XY}\Sigma(f_R\Delta C_0^R)$, where $\theta = \chi_1 + \Delta\chi_1$, as described in the text (Eq. (3)). Doubly bonded carbonyl or carboxyl oxygen substituents receive two counts, whereas increments for singly bonded hydroxyl or carboxyl oxygen count once.

the initially inner substituent C^γ becoming part of the fundamental coupling path requires its associated increment be removed, resulting in counts (as shown on the first row of each block on the left) all being decremented by one unit (on the right) for residues possessing C^γ.

Substituents to both the terminal site X (H^α, N′ or C′) and the central atom C^α remain constant within both these series. However, cycling the active coupling partner within each series requires first, if X equals H^α, no terminal increment, second, if X equals N′, the inclusion of a single increment to reflect the adjacent C'_{i-1} present in the preceding amino-acid residue and, finally, if X equals C′, the inclusion of two terminal substituents for the backbone atoms O'_i and N'_{i+1} from current and subsequent residues, respectively. Patterns at the C^α position change accordingly. The sets of three fundamental Karplus coefficients established separately for each combination of the coupled nuclei X and Y already account for altering the constituent four-atom fragments indicated in the center of each panel of Fig. 1.

Heavy-atom types are substituted as necessary for the default atom type H at central or inner (C^α or C^β) and terminal or outer (N′, C′, or C^γ) positions. Departing from, for example, ethane as a template for a basic four-atom $^3J_{H\alpha,H\beta}$ moiety, the majority of the amino-acid types involve one inner N and two inner C increments reflecting N′ and C′ bonded to C^α, and C^γ bonded to C^β (Fig. 1). Terminal substituents are obsolete with $^3J_{H,H}$ couplings, giving a respective void block in Table 1. When departing from propane as a template, $^3J_{C',H\beta}$ already takes into account the C′ atom and only a single inner C increment is normally required due to the presence of C^γ, as well as one N and two O outer increments to reflect singly-bonded N'_{i+1} of the subsequent residue and doubly-bonded carbonyl O'_i of the current residue, respectively.

Identical substituent patterns shared among all amino acids result in strings of equal counts within Table 1 rows. For example, all χ_1 topologies comprise both N′ and C′ as *inner first-sphere* substituents to C^α, giving rise to two underlying runs of '1' within blocks $^3J_{H\alpha,H\beta}$ and $^3J_{H\alpha,C\gamma}$. When the N′ substituent turns into an active coupling partner in blocks $^3J_{N',H\beta}$ and $^3J_{N',C\gamma}$, the associated counts are eliminated and only those of C′ retained. Conversely, only counts of N′ are retained in blocks $^3J_{C',H\beta}$ and $^3J_{C',C\gamma}$.

As for the *outer first-sphere* substituents, the peptide linkages manifest in a count for the carbonyl carbon of the preceding residue in blocks $^3J_{N',H\beta}$ and $^3J_{N',C\gamma}$, while absent in other blocks. Blocks $^3J_{C',H\beta}$ and $^3J_{C',C\gamma}$ display two rows of increments for carbonyl oxygen in the current residue, and amide nitrogen in the subsequent one. Twofold increments for carbonyl oxygen are to reflect the double-bond character, whereas single counts are being inserted for side-chain hydroxyl groups in Ser, Thr, and Tyr.

Although all atoms in fragments X–C^α–C^β–Y form part of the same amino acid, some substituents to both X and C^α sites extend into adjacent preceding or subsequent

residues. Given the high similarity of $J(\chi_1)$ coupling topologies with regard to their *backbone directed* portion, the models here predict *first-sphere* substituent effects not to depend on the actual polypeptide sequence.

2.3. Second-sphere substituent patterns

Including increments for second-sphere substituents enhances the amino-acid type differentiation. Construction of the second-sphere increments follows the procedure outlined for the first sphere, except that substituents will be two bonds away, considering the shortest path, from the constituent four-atom fragment.

Effects are more pronounced for χ_1 than for ϕ. Considering the *backbone directed* portion in χ_1, although the terminal site $X = H^{\alpha}$ does not give rise to outer substituents, all three peptide-link heavy atoms appear as three rows of *inner second-sphere* substituents to H^{α} in blocks $^3J_{H\alpha,H\beta}$ and $^3J_{H\alpha,C\gamma}$. Again, mutual elimination of the other substituents in blocks involving active N′ and C′ coupling partners is a consequence of their being accounted for as *outer first-sphere* substituents. Finally, entries for *outer second-sphere substituents* reflect C^{α} and O′ of preceding residues in blocks involving $X = N'$, and C^{α} of subsequent residues in blocks involving $X = C'$.

The *side-chain directed* portion exhibits a characteristic column pattern that is most prominent for χ_1 related couplings involving C^{γ} as an active coupling partner. As $^3J_{H\alpha,H\beta}$ couplings do not involve terminal substituents at all, non-zero entries in corresponding places for $^3J_{H\alpha,C\gamma}$ immediately reveal the topological effects brought about by switching between H^{β} and C^{γ} (left vs. right blocks in Table 1). This pattern overlays row entries in the remaining two blocks involving C^{γ}. Somewhat concealed, a similar recurrent inner-substituent pattern affects H^{β} related couplings. It is observed that patterns due to *outer first-sphere* substituents in C^{γ} related blocks replicate as *inner second-sphere* substituents in H^{β} related blocks.

The proline 5-ring system deserves special attention as the model duplicates increments to some endocyclic couplings. Second-sphere substituents shared between both X and Y site were counted twice, based on the rationale that they would if this were an enlarged 6-ring system. This concerns C^{δ} and C^{γ} in χ_1 and ϕ applications, respectively.

3. Results

Six different ^3J-coupling models were self-consistently fitted to sets of coupling constants measured in the protein flavodoxin, related either to χ_1 or to ϕ torsions. As summarized in Table 2, the combinations exclude (M_0) or include (M_1) angular Gaussian random motion about the mean torsion angles [18], disregard (S_0) or take into account substituent effects up to the first (S_1) or second sphere (S_2).

The non-linear least-squares optimizations yielded, apart from the sought torsion angles, coupling increments attributed to each type of substituent atom encountered. Analyses at the S_2 level comprised 12 χ_1 related (Table 3) or 10 ϕ related (Table 4) substituent-specific increment parameters, ΔC_0. Combination of these with the respective sets of three fundamental Karplus coefficients proved sufficient to account for first- and second-sphere substituent

Table 2
Comparison of ^3J coupling models fitted to *D. vulgaris* flavodoxin data related to torsions χ_1 and ϕ[a]

	M_0S_0	M_0S_1	M_0S_2	M_1S_0	M_1S_1	M_1S_2
Number of observables ($n = 763$)	Side-chain torsion angles χ_1 (112)					
$RMSD_J$ (Hz)	0.54	0.47	0.47	0.47	0.40	**0.39**
$RMSD_\theta$ (°)	48.3	39.0	41.9	45.9	39.9	**42.4**
Number of adjustables (p)	130	136	142	242	248	**254**
Degrees of freedom (n–p)	633	627	621	521	515	**509**
Normalized fit error ε_J^2 ($\sigma_J = 0.50$)	876.8	684.7	673.0	664.1	486.5	**469.7**
Abs. significance Q (%)	0.0	5.5	7.3	0.0	81.1	**89.3**
Number of observables ($n = 776$)	Main-chain torsion angles ϕ (135)					
$RMSD_J$ (Hz)	0.36	0.36	0.36	0.31	**0.31**	0.31
$RMSD_\theta$ (°)	6.8	6.8	6.8	8.7	**7.7**	7.9
Number of adjustables (p)	153	157	163	288	**292**	298
Degrees of freedom (n–p)	623	619	613	488	**484**	478
Normalized fit error ε_J^2 ($\sigma_J = 0.40$)	640.1	628.9	624.0	472.7	**463.7**	457.9
Abs. significance Q (%)	30.9	37.2	34.9	68.2	**72.8**	71.6

[a] *Notes:* Self-consistent J-coupling modelling refers to a minimum 4 coupling constants per residue, with Gaussian-random angle libration disabled (M_0) or enabled (M_1), using 6×3 fundamental Karplus coefficients, either ignoring substituent effects (S_0) or including component-coupling increments from first-sphere substituents (S_1) or, in addition, from the second sphere (S_2). $RMSD_\theta$ refer to comparison torsion angles from X-ray coordinates of the wild-type *D. vulgaris* flavodoxin-FMN complex (oxidized) resolved at 0.17 nm [12]. However, $RMSD_\theta$ is *not* a measure of correctness, and possible reasons for angle discrepancies between NMR and X-ray derived structures were discussed previously [8]. Absolute significances, $Q = 1 - P$, in the context of χ^2-statistics derive from tests of the incomplete Gamma probability-distribution function [23], where the probability, $P = \Gamma(\chi^2/2, \nu/2)$, is the likelihood that the fitted parameter set is a chance incidence. As a rule of thumb, $P \approx 0.5$ if the *normalized* χ^2 error (denoted ε_J^2 in the present work to avoid confusion with torsion-angle notations) approaches the number of degrees of freedom, $\nu = n - p$, in the model. Best fitting models are indicated in bold.

Table 3
Self-consistent component couplings ΔC_0^R (Hz) related to polypeptide χ_1 substituents[a]

Substituent						ΔC_0^R	
Type	Sphere	Site (X)	R = H →	Site (Y)	R = H →	M_0S_2 (rigid)	M_1S_2 (Gauss)
Inner/central	1st	C^α	C'	C^β	C^γ	−0.37	−0.70
		C^α	N'	—	—	n/a	n/a
		—	—	C^β	O^γ/O^η	−1.21	**−1.60**
		—	—	C^β	S^γ	−1.16	**−1.43**
	2nd	C^α	C'_-/C^δ	C^β	C^δ	+0.08	**−0.04**
		C^α	N'_+	C^β	N^δ	−0.03	**+0.03**
		C^α	O'	C^β	O^δ	+0.09	**−0.02**
		—	—	C^β	S^δ	n/a	n/a
Outer/terminal	1st	N'	C'_-/C^δ	C^γ	C^δ	−0.25	**−0.31**
		C'	N'_+	C^γ	N^δ	+2.21	**+1.58**
		C'	O'	C^γ	O^δ	+0.44	**+0.54**
		—	—	C^γ	S^δ	n/a	n/a
	2nd	N'/C'	C^α_-/C^α_+	C^γ	C^ε	+0.21	**+0.22**
		—	—	C^γ	N^ε	−0.16	**−0.02**
		N'/C'	O'_-	C^γ	O^ε/O^η	−0.36	**−0.41**
		—	—	C^γ	S^ε	n/a	n/a

[a] *Notes:* Incremental substituent effects, ΔC_0^R, from substituting heavy atoms for the default type H bonded to central (C^α, C^β) or terminal (N', C', C^γ) positions. Atoms subscripted '−' and '+' form part of residues '$i-1$' and '$i+1$', respectively, preceding and following the current residue i for which no subscript is being used. Shared among all amino-acid and *J*-coupling types using appropriate weights, the increments ΔC_0^R augment the fundamental Karplus coefficients C_0 obtained in the same analysis. Absence and presence of angular motion in the model is indicated. Bold values were used to compile amino-acid specific Karplus curves given in Table 5. Calculations are based on data collected for *D. vulgaris* flavodoxin, where n/a indicates that the respective increment is not encountered in this amino-acid sequence. For example, no methionine data were available to study the effect of S^δ substituents. N' bonded as inner first-sphere substituent to C^α does not involve a coupling increment as its uniform effect in all amino acids is reflected in the fundamental coefficients already.

Table 4
Self-consistent component couplings ΔC_0 (Hz) related to polypeptide ϕ substituents[a]

Substituent						ΔC_0^R	
Type	Sphere	Site (X)	R = H →	Site (Y)	R = H →	M_0S_2 (rigid)	M_1S_2 (Gauss)
Inner/central	1st	N'	C'_-	C^α	$C'_-/C^\beta/C^\delta$	+0.23	**+0.61**
		—	—	—	—	n/a	n/a
		—	—	—	—	n/a	n/a
		—	—	—	—	n/a	n/a
	2nd	N'	$C^\alpha_-/(C^\gamma)$	C^α	C^γ	−0.02	**−0.05**
		—	—	C^α	N'_+	n/a	n/a
		N'	O'_-	C^α	O'/O^γ	±0.00	**−0.32**
		—	—	C^α	S^γ	+0.43	**−0.47**
Outer/terminal	1st	C'_-/C'	C^α_-/C^α_+	C^β	$C^\gamma/(C^\varepsilon)$	−0.26	**−0.22**
		C'	N'_+	—	—	n/a	n/a
		C'_-/C'	O'_-/O'	C^β	O^γ	−0.02	**+0.26**
		—	—	C^β	S^γ	−0.44	**+0.22**
	2nd	C'_-/C'	$C^\beta_-/C^\beta/(C^\delta_+)$	C^β	C^δ	−0.06	**−0.13**
		C'_-/C'	N'_-	C^β	N^δ	−0.04	**−0.07**
		—	—	C^β	O^δ	−0.06	**−0.13**
		—	—	C^β	S^δ	n/a	n/a

[a] *Notes:* Incremental substituent effects, ΔC_0^R, from substituting heavy atoms for the default type H bonded to central (N', C^α) or terminal (C'_-, C', C^β) positions. Some inner first-sphere substituent combinations do not occur, such as nitrogen (other than the one forming the coupling fragment), oxygen or sulfur bonded to N' or C^α. N'_+ bonded as inner second-sphere substituent to C^α, and as outer first-sphere substituent to C', does not involve coupling increments as its uniform effect in all amino acids is reflected in the fundamental coefficients already. No methionine data were available to study the effect of S^δ substituents. Bold values were used to compile amino-acid specific Karplus curves given in Table 7. Other details apply as given in Table 3 footnote.

effects on χ_1 and ϕ torsion-angle dependencies of ^3J couplings in the common amino-acid types. Fig. 2 demonstrates how the *J*-coupling discrepancies between calculated and experimental values improve with the more sophisticated models. In line with previous experience [24], the remaining discrepancies showed no correlation with the actual dihedral-angle value.

3.1. Amino-acid χ_1 torsion related couplings

The presence of two tetrahedral entities, C^α and C^β, in the χ_1 coupling path creates up to nine possible ^3J pair interactions. Diastereospecific considerations notwithstanding [6], χ_1 torsions across all amino-acid side chains involve only six different *J*-coupling types, namely $^3J_{H\alpha,H\beta}$,

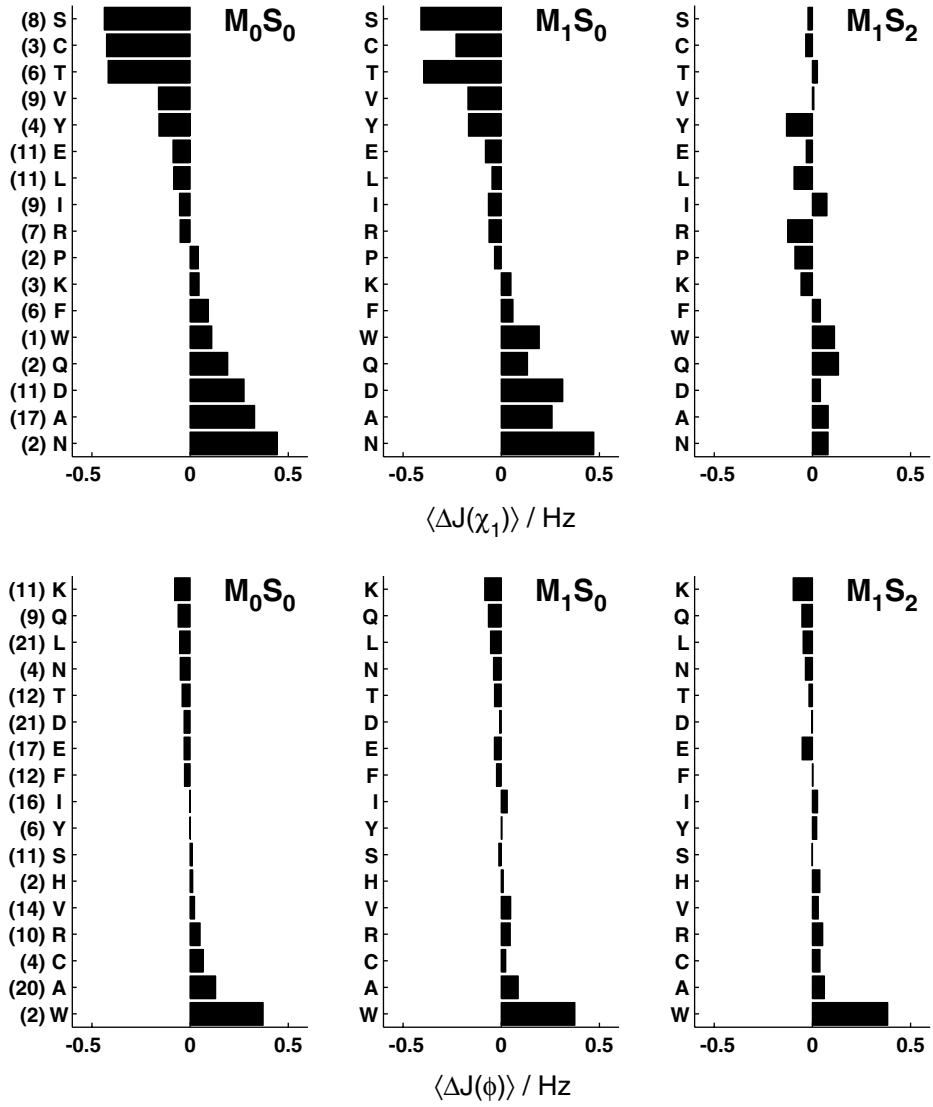

Fig. 2. J discrepancies, $\Delta J = J_{expt} - J_{calc}$, averaged over all contributing coupling types, separately for each amino-acid type (one-letter residue code, frequencies in parentheses). Results from studies of χ_1 (top) and ϕ (bottom) are given for (M_0S_0) the conventional rigid-torsion model, (M_1S_0) including Gaussian-type angular motion, and (M_1S_2) substituent effects up to the second sphere in addition. Fit discrepancies shown are aggregate average deviations from predicted values and, thus, do not reflect the spread of error within each amino-acid class or J-coupling type. Regardless of the model applied, residual fit errors in the ϕ analysis are nearly all within the experimental error of J-coupling determination. The systematically larger experimental – or smaller calculated – values for tryptophan may be a consequence of insufficient sampling (two occurrences only) or due to limitations to the model with respect to aromatic moieties.

$^3J_{N',H\beta}$, $^3J_{C',H\beta}$, $^3J_{H\alpha,C\gamma}$, $^3J_{N',C\gamma}$ and $^3J_{C',C\gamma}$. The corresponding torsion-angle phase increments are all multiples of 120° as the intervening coupling path involves a sp^3–sp^3 configuration.

Incremental substituent effects related to protein side-chain torsions χ_1 were calculated from an experimental set of 746 ^3J coupling constants collected for *D. vulgaris* flavodoxin [8]. Coefficients were eventually derived including data available for alanine residues duplicated, as previously described, raising to 763 the total number of reference points used in the simultaneous optimization of all χ_1 angles. Assuming standard errors of 0.50 Hz in J and, for the moment, leaving aside the improved S_2 models, calculations for the 95 χ_1 torsions (excluding alanines) at the S_1 level essentially reproduce the results given in [8].

Substituent effects on $^3J(\chi_1)$ are statistically highly significant. In order for significance to be accomplished, a more complex model requires the normalized residual fit error drop by more units than the increase in the number of fit parameters. When proceeding from S_0 toward S_1 models, introducing only 6 increment parameters attributed to first-sphere effects lowered the fit error by typically 200 normalized units (Table 2). Another six parameters for second-sphere effects gave an additional improvement of 20 units with S_2 models. Similarly, allowing for torsion-angle libration by adding 112 Gaussian spread parameters also decreased the residual by approximately 200 normalized units, placing higher statistical significance on M_1 models than on M_0.

It is observed that χ_1-related substituent effects on J significantly exceed the experimental precision of determining J, underlining the importance to a reliable χ_1 analysis of taking longer-ranging substituent effects into account. Consensus Karplus curves for $^3J_{H\alpha,H\beta}$, $^3J_{N',H\beta}$, $^3J_{C',H\beta}$, $^3J_{H\alpha,C\gamma}$, $^3J_{N',C\gamma}$ and $^3J_{C',C\gamma}$ shown in Table 5 and Fig. 3 differ from those previously obtained for first-sphere substituents only [8], on average, by 0.80, 0.17, 0.26, 0.22, 0.15, and 0.21 Hz [25]. Rotationally averaged values of ^3J, as given by the angle-independent coefficients C_0, are virtually identical between this study and the previous, meaning that the overall coupling magnitudes were reproduced in both fits. Modulation amplitudes, however, as signified by coefficients C_2, became larger with the new higher-order substituent model. In particular, *trans* $^3J_{H\alpha,H\beta}$ couplings rose by 0.6 Hz as a result of the improved coefficients. This points to a neater separation of mobility and substituent effects, as anticipated previously [8].

Differences between the largest and smallest C_0 coefficients (Table 5) of 2.31, 0.74 and 1.15 Hz for couplings involving H^β, on the one hand, and 2.14, 0.68 and 1.07 Hz for those involving C^γ, on the other hand, are also seen as residue-type dependent spreads within manifolds of Karplus curves in the individual panels of Fig. 3. These spreads parallel the scaling factors γ_{XY} (Eq. (3)), equaling 1.0000, 0.3183 and 0.5014, and reflect the presence of the coupling partner $X = H^\alpha$, N' and C', respectively.

Both $^3J_{H\alpha,H\beta}$ and $^3J_{H\alpha,C\gamma}$ show similar residue-type sensitivity, and likewise other pairs too. Seemingly, the cumulative substituent effects on all J couplings involving C^γ effectively make up for any scaling from its lower gyromagnetic ratio compared to H^β.

The side chains in Thr, Ser and Cys provide only limited sets of up to six χ_1 related coupling constants as oxygen and sulfur substituents are NMR inactive. Threonine consistently and invariably exhibits the smallest $^3J(\chi_1)$ values across all coupling types, due chiefly to the combined effects of both methyl and hydroxyl groups present. Coupling constants involving H^β in Ser and Cys residues are likewise small. This contrasts with notoriously above-average coupling constants, especially those involving actively coupled C^γ in residues Asn and Asp, and likely His, too, attributed to electron-withdrawing effects exerted by oxygen and nitrogen substituents in terminal positions, as opposed to the electron-donating oxygen situated on the intervening path in Thr.

Linear, unbranched side-chains exhibit coupling constants close to consensus values derived from the complete ensemble of amino acids in the fit (Table 5). The presence of aliphatic C–(CH)–C branches in the Val and Ile side-chain topologies, which means that $C^{\gamma 2}$ replaces one of the methylene H^β atoms, appears to lower the values of all coupling types, consistent with quantum mechanical calculations on substituted hydrocarbons [26]. Conversely, alanine gives rise to the largest coupling constants (only those involving H^β are available) as carbon-framework branches are not present at all. This branch effect is mirrored also in Thr *vs.* Ser couplings.

Couplings in aromatic amino acids behave similarly to those in simpler linear chains, pointing to a vanishing influence of remote carbon substituents. However, experimental data for tyrosine C^γ related couplings appear to be slightly smaller than those in phenylalanine, which is why an increment was included to reflect the presence of the remote oxygen in a homoallylic second-sphere position.

The fact that Asx and Glx residues form the extremes on the coupling-value scale involving C^γ is consistent with the observation of so-called α and β effects.

3.2. Amino-acid ϕ torsion related couplings

The polypeptide backbone torsion ϕ also gives rise to six different ^3J-coupling types, namely $^3J_{HN,H\alpha}$, $^3J_{HN,C'}$, $^3J_{HN,C\beta}$, $^3J_{C',H\alpha}$, $^3J_{C',C'}$ and $^3J_{C',C\beta}$. As the intervening coupling path in X–N′–C^α–Y consists of a sp^2–sp^3 configuration, torsion-angle phase increments are unique multiples of 60°, with no redundance present like in the case of χ_1.

Extending the initial investigation into self-consistent Karplus parameters related to the main-chain torsion ϕ [7], this work includes amino-acid specific effects on $^3J(\phi)$ from both first-sphere and second-sphere substituents (Table 6).

Reference data related to flavodoxin backbone torsions ϕ comprised 776 experimental ^3J coupling constants, including those reported previously [7], augmented by additional data from 13 glycine residues. Proline lacks all ϕ-related ^3J couplings involving H^N and was therefore excluded from the fit, as the inclusion threshold was a minimum four coupling constraints per residue, taking the total number of torsions fitted to 135. Experimental standard errors of 0.40 Hz were applied to ϕ-related coupling constants, tighter than for those related to χ_1 in order to reflect the typically higher experimental precision achieved for these J coupling types.

Normalized fit errors decreased by approximately 160 units on the inclusion in the ϕ fit of 135 additional Gaussian spread parameters of approximately 20° to account for torsion-angle libration (Table 2), indicating that M_1 models perform statistically better than M_0, though perhaps not significant in practice (Fig. 2). Average residual fit violations in J for models M_0 and M_1 were 0.36 Hz and 0.31 Hz, respectively, regardless of the level of substituent effects considered. Torsion-angle discrepancies with X-ray data [12], averaging generally less than ±9°, were well within thermal angular libration range and therefore negligible.

Compared with previous parametrisations ignoring substituent effects [7], the new consensus Karplus coefficients at the second-sphere level (Table 7) anticipate average differential effects [25] of 0.46, 0.49, 0.18, 0.57, 0.15, and 0.21 Hz, respectively, for the ϕ related couplings $^3J_{HN,H\alpha}$, $^3J_{HN,C'}$, $^3J_{HN,C\beta}$, $^3J_{C',H\alpha}$, $^3J_{C',C'}$ and $^3J_{C',C\beta}$. Noticeable effects exceeding experimental error margins are thus expected for three out of the six coupling types. Amino-acid

Table 5
Amino-acid χ_1 related Karplus coefficients self-consistently inferred from flavodoxin ^3J couplings[a]

J type	χ_1 substituent pattern	C_0	C_1	C_2	A	B	C	J_{trans}	J_{gauche}
	Fundamental	*7.51*					*3.13*	*13.08*	*4.73*
	Ala	6.77					2.38	12.33	3.98
	Asn	6.06					1.68	11.63	3.28
	Arg, Glx, Leu, Lys, Met	6.02					1.63	11.58	3.24
	Asp, His	6.01					1.62	11.57	3.23
	Pro	5.98					1.59	11.54	3.19
	Phe, Tyr, Trp	5.93					1.54	11.50	3.15
$^3J(H^{\alpha},H^{\beta})$	**Consensus**	**5.85**	**−1.18**	**4.39**	**8.77**	**−1.18**	**1.46**	**11.41**	**3.07**
	Val	5.36					0.97	10.92	2.58
	Cys	5.34					0.95	10.90	2.55
	Ile	5.31					0.93	10.88	2.53
	Ser	5.16					0.78	10.73	2.38
	Thr	4.46					0.07	10.02	1.68
	Fundamental	*3.03*					*1.71*	*5.09*	*2.00*
	Ala	2.52					1.20	4.58	1.48
	Asn, Met	2.29					0.98	4.36	1.26
	Arg, Asp, Glx, His, Leu, Lys	2.28					0.96	4.34	1.25
	Phe, Pro, Tyr, Trp	2.25					0.93	4.31	1.22
$^3J(N',H^{\beta})$	**Consensus**	**2.22**	**−0.75**	**1.32**	**2.63**	**−0.75**	**0.91**	**4.29**	**1.19**
	Val	2.07					0.75	4.13	1.04
	Cys	2.06					0.74	4.12	1.03
	Ile	2.05					0.74	4.12	1.02
	Ser	2.01					0.69	4.07	0.97
	Thr	1.78					0.47	3.84	0.75
	Ala	3.77					1.50	7.63	1.85
	Asn	3.42					1.15	7.28	1.49
	Arg, Asp, Glx, His, Leu, Lys, Met	3.40					1.13	7.25	1.47
	Pro	3.38					1.11	7.23	1.45
	Phe, Tyr, Trp	3.36					1.08	7.21	1.43
$^3J(C',H^{\beta})$	**Consensus**	**3.32**	**−1.58**	**2.27**	**4.54**	**−1.58**	**1.04**	**7.17**	**1.39**
	Val	3.07					0.80	6.92	1.14
	Cys	3.06					0.79	6.91	1.13
	Ile	3.05					0.77	6.90	1.12
	Ser	2.97					0.70	6.82	1.04
	Thr	2.62					0.35	6.47	0.69
	Fundamental	*2.35*					*0.08*	*6.21*	*0.42*
	Asn	5.00					2.09	8.83	3.08
	Asp	4.48					1.57	8.31	2.56
	His	4.25					1.35	8.08	2.33
	Fundamental	*4.04*					*1.14*	*7.87*	*2.12*
	Met	3.78					0.87	7.61	1.86
	Lys	3.62					0.72	7.46	1.70
	Phe	3.53					0.63	7.37	1.62
	Trp	3.52					0.62	7.36	1.61
	Arg	3.50					0.60	7.33	1.58
	Pro	3.49					0.59	7.32	1.57
$^3J(H^{\alpha},C^{\gamma})$	**Consensus**	**3.43**	**−0.93**	**2.90**	**5.80**	**−0.93**	**0.53**	**7.26**	**1.51**
	Leu	3.35					0.45	7.19	1.44
	Tyr	3.33					0.43	7.16	1.41
	Val	3.31					0.41	7.15	1.39
	Ile	3.16					0.25	6.99	1.24
	Gln	3.09					0.19	6.92	1.17
	Glu	2.89					−0.01	6.73	0.98
	Thr	2.86					−0.04	6.69	0.94
	Asn	1.53					0.74	2.76	0.92
	Asp	1.37					0.58	2.59	0.75
	Fundamental	*1.36*					*0.57*	*2.59*	*0.75*
	His	1.29					0.50	2.52	0.68
	Met	1.14					0.35	2.37	0.53
	Lys	1.09					0.30	2.32	0.48
	Phe	1.07					0.28	2.29	0.45

(*continued on next page*)

Table 5 (*continued*)

J type	χ_1 substituent pattern	C_0	C_1	C_2	A	B	C	J_{trans}	J_{gauche}
	Trp	1.06					0.27	2.29	0.45
	Arg	1.05					0.26	2.28	0.44
$^3J(N',C^\gamma)$	**Consensus**	**1.03**	**−0.44**	**0.79**	**1.58**	**−0.44**	**0.24**	**2.26**	**0.42**
	Leu	1.01					0.22	2.23	0.40
	Pro	1.01					0.22	2.23	0.40
	Tyr	1.00					0.21	2.23	0.39
	Val	0.99					0.20	2.22	0.38
	Ile	0.95					0.16	2.17	0.33
	Gln	0.92					0.13	2.15	0.31
	Glu	0.86					0.07	2.09	0.25
	Thr	0.85					0.06	2.08	0.24
	Asn	2.50					1.14	4.71	1.39
	Asp	2.24					0.88	4.44	1.13
	His	2.12					0.77	4.33	1.02
	Met	1.88					0.53	4.09	0.78
	Lys	1.81					0.45	4.02	0.70
	Phe, Trp	1.76					0.41	3.97	0.66
	Arg	1.75					0.39	3.96	0.64
	Pro	1.74					0.39	3.95	0.64
$^3J(C',C^\gamma)$	**Consensus**	**1.71**	**−0.86**	**1.35**	**2.71**	**−0.86**	**0.36**	**3.92**	**0.61**
	Leu	1.67					0.32	3.88	0.57
	Tyr	1.66					0.31	3.87	0.56
	Val	1.65					0.30	3.86	0.55
	Ile	1.57					0.22	3.78	0.47
	Gln	1.54					0.19	3.75	0.44
	Glu	1.44					0.09	3.65	0.34
	Thr	1.43					0.07	3.63	0.32
	Fundamental	*1.12*					*−0.24*	*3.32*	*0.01*

[a] *Notes*: Coefficients (in Hz) for direct use with the Fourier-type representation of the Karplus equation, $^3J(\theta) = C_0 + C_1 \cos\theta + C_2 \cos 2\theta$,where $\theta = \chi_1 + \Delta\chi_1$, with $\Delta\chi_1$ depending on the actual diastereospecific positions of the coupled nuclei. Alternative coefficients are given for the power-series representation of the Karplus equation, $^3J(\theta) = A\cos^2\theta + B\cos\theta + C$. Coefficients C_1 and C_2 (or A and B) use the amino-acid independent *consensus* values. Coefficients C_0 (or C) given already comprise incremental effects according to Eq. (3), i.e., amino-acid independent *fundamental* coefficients (for a fully hydrogenated fragment $X–C^\alpha–C^\beta–Y$) combined with substituent related increments (Table 3), weighted depending on amino-acid topology (Table 1), to form amino-acid specific coefficients. *Consensus* coefficients are averages over the 112 residues included in the fit, weighted by their fractional type occurrence: Ala (17), Arg (7), Asn (2), Asp (11), Cys (3), Gln (2), Glu (11), His (0), Ile (9), Leu (11), Lys (3), Met (0), Phe (6), Pro (2), Ser (8), Thr (6), Trp (1), Tyr (4), Val (9). Results derive from calculations including up to second-sphere substituent effects and Gaussian-random fluctuation (M_1S_2 model).

specific spreads in $^3J(\phi)$ derived from the largest and smallest values of C_0 for each coupling type (Fig. 4) were 0.61, 0.31, 0.44, 0.58, 0.28, and 0.22 Hz, partly reflecting the γ weighting applied (Eq. (3)) of 1.0000, 0.5014 and 0.2515 for H–H, H–C and C–C connectivities, respectively.

Although substituent influences appear to be substantially smaller on main-chain than on side-chain related couplings, performance of the improved model is nevertheless statistically significant as the addition of only 10 fit parameters, ΔC_0, diminishes the total fit residual, ε^2, by about 15 normalized units in the S_2 model. Even S_1 models, which attribute 4 increment parameters to first-sphere effects ensuing 10 normalized units improvement in the fit error (Table 2), are deemed more effective than S_0 models which disregard substituent effects entirely. By this measure, the simpler S_1 model turns out to be marginally more significant, if only on the basis of the present flavodoxin data. Fig. 2 shows the residual J violations broken down by residue type.

When site Y in a ϕ-related coupling involves H^α or C′, the branched side-chains Val and Ile show marginally larger J values than consensus, while oxygen and sulfur bearing amino acids Ser, Thr and Cys figure at the bottom of the value range, penultimate only to Gly which is notorious for the smallest J values. This residue-type dependent order is roughly reversed for $^3J_{HN,C\beta}$ and $^3J_{C',C\beta}$, both involving the side-chain carbon C^β as an active coupling partner.

Trivially, glycine does not give rise to couplings in which $Y = C^\beta$. Yet, from Gly to Ala the average $^3J_{HN,H\alpha}$ coupling increases by 0.61 Hz, reflecting the effect of substituting the side-chain carbon C^β for the H^α proton.

Dependencies of $^3J(\phi)$ on amino-acid topology are resolved only for those couplings involving the terminal site C^β as this is the only coupling path to sense variation in second-sphere side-chain topology, thus discriminating the different side-chain types. Residue-type dependent spreads are therefore widest for $^3J_{HN,C\beta}$, though generally more narrow than for the χ_1 related Karplus curves.

Although not all of the 20 amino-acid types were represented in the flavodoxin dataset, the modular concept of the model allows predictions to be made for yet inaccessible data. Proline is predicted consistently to rank top in the $X = C'$ coupling sets, especially $^3J_{C',H\alpha}$ and $^3J_{C',C'}$. Proline ϕ-related couplings are difficult to measure, which is why

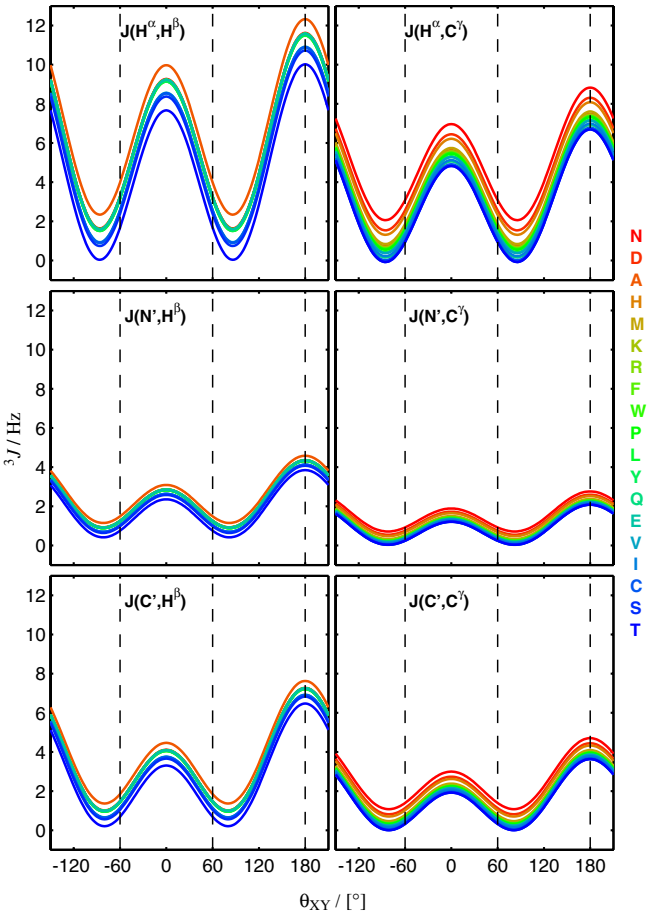

Fig. 3. Amino-acid type specific Karplus curves for side-chain $^3J(\chi_1)$ couplings. Angle-independent offsets reflect topological variation in the different amino acids. The self-consistent model applied, M_1S_2, included Gaussian-type angular motion and second-sphere substituent effects. Vertical dashed lines indicate *trans* and *gauche* orientations of the dihedral angle subtended, signposting ^3J values used for determining staggered-rotamer distributions. Matching the rainbow-coloured curves, one-letter codes on the right rank the amino-acid types by their average coupling magnitude as seen in the flavodoxin dataset.

experimental values were not available for comparison at this time.

4. Discussion

J-coupling values calculated from self-consistent, amino-acid specific Karplus coefficients matched the experimental constraints dramatically better than using parameter sets not differentiating between the various substituent patterns. Taking the nature and position of substituents into account, and their possible impact on the ^3J-coupling constants, previously proved to be critical to the reliable analysis of polypeptide side-chain χ_1 torsion angles [8], whereas this did not appear as crucial in the analysis of ϕ torsions [7].

Whether topological effects were included or excluded, coupling constants back-calculated for the rigid ϕ torsion model (M_0 level) converged with experimental ones within measurement precision ($RMSD_J = 0.36$ Hz in Table 2). Allowing for moderate Gaussian angular libration reduced the discrepancy between predicted and experimental coupling constants to 0.31 Hz. Corroborating common experience, substituent effects seem to have little impact on backbone related $^3J(\phi)$ couplings.

Protein side-chain related $^3J(\chi_1)$ couplings, however, are equally sensitive to both angular mobility and substituent effects. In the case of χ_1, $RMSD_J$ decreased on inclusion of either mobility or topology effects by equal proportions, to 0.47 Hz from 0.54 Hz. Both effects combined reduced the fit residual further to 0.39 Hz.

The empirical component coupling model here proposed rationalizes as to why substituent effects on backbone $^3J(\phi)$ couplings appear negligible in practice. For variation in increment counts, both first- and second-sphere, is virtually absent among the set of amino acids, and is confined to couplings involving the C^{β} branch only. Analyses of peptide and protein ϕ-related coupling constants have not traditionally considered such effects either. Susceptible to topological variation are predominantly those coupling types that involve actively coupled nuclei exposed to the side-chain branch, i.e., C^{γ} in the case of χ_1 and C^{β} in the case of ϕ. As there are three such coupling types related to χ_1, as opposed to only two in the case of ϕ, qualitatively stronger ^3J coupling dependencies on amino-acid type are anticipated for side-chain than for backbone.

These insights lend to believe that the greater variety of functional groups around the more remote side-chain torsions, such as χ_2, should produce significant measurable effects.

The component coupling model also predicts substituent influences on amino-acid backbone ψ-related ^3J couplings to be even smaller than on those related to ϕ. There are other principal difficulties, however, in determining quantitative ψ torsion angles. First, as ψ torsions relate to only three ^3J-coupling constants, a shortage of experimental data prevents resolving all ambiguities in the Karplus relation. Second, ψ-related ^3J coupling constants are notoriously small and barely exceed 1 Hz [27], owing to the involvement of low-gyromagnetic ratio nitrogen in all coupling paths.

Donders et al. [28] addressed the possibility of non-linear effects arising from compiling multiple substituent related electronegativities or component couplings. Given their use of synthetic, calculated rather than experimental coupling constants and the comparatively small magnitude of effects, well below measurement precision in proteins, such effects were considered negligible in the present investigation.

Terminal substituents in a polypeptide backbone coupling topology may form part of adjacent amino-acid residues, making the analysis potentially dependent on the explicit amino-acid sequence, although effects over two bonds are unlikely to sense the type of the adjacent side chain. Effects from substituents more remote than two bonds are considered negligible on the grounds that second-sphere substituent effects already diminish rapidly with their increased distance from the focal atoms. First-sphere substituents, those directly attached to either coupled or

Table 6
Substituent counts m for amino-acid specific ^3J coupling dependences on the torsion ϕ[a]

Type	Sphere	R	Gly	Thr	Ile	Val	His	Trp	Tyr	Phe	Leu	Met	Gln	Glu	Lys	Arg	Asn	Asp	Ala	Ser	Pro	Cys	Gly	Thr	Ile	Val	His	Trp	Tyr	Phe	Leu	Met	Gln	Glu	Lys	Arg	Asn	Asp	Ala	Ser	Pro	Cys
			$^3J_{HN,H\alpha}$																				$^3J_{C',H\alpha}$																						
Inner/central	1st	C	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	2nd	C	1	2	3	3	1	2	2	2	2	2	2	2	2	2	2	2	1	1	2	1	—	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	—	—	1	—
		N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	4	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	4	4	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
Outer/terminal	1st	C	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	2nd	C	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
			$^3J_{HN,C'}$																				$^3J_{C',C'}$																						
Inner/central	1st	C	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	—	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	2nd	C	1	2	3	3	1	2	2	2	2	2	2	2	2	2	2	2	1	1	2	1	—	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	—	—	2	—
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		O	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
Outer/terminal	1st	C	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	2nd	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

			$^3J_{HN,C\beta}$																				$^3J_{C',C\beta}$																			
Inner/central	1st	C	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	2nd	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		O	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Outer/terminal	1st	C	—	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	—	—	1	—	1	2	3	3	1	2	2	2	2	2	2	2	2	2	2	2	1	1	2	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		O	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2
		S	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1
	2nd	C	—	—	1	—	2	3	3	3	2	—	1	1	1	1	—	—	—	—	1	—	1	1	2	1	3	4	4	4	3	1	2	2	2	2	1	1	1	1	2	1
		N	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		O	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	3	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	3	—	—	—	—
		S	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—

[a] *Notes*: Details apply as given in the Table 1 footnote, except that $\theta = \phi + \Delta\phi$ in the extended Karplus equation $^3J(\theta) = C_0 + C_1 \cos\theta + C_2 \cos 2\theta + \gamma_{XY}\Sigma(f_R \Delta C_0^R)$.

intervening atoms, are generally more influential than those in the second sphere.

4.1. Application of the χ_1 related coefficients to other proteins

Of interest is the performance of the amino-acid specific Karplus parametrizations obtained when applied to coupling data from proteins other than flavodoxin. As Fig. 2 demonstrates, residual fit errors in the amino-acid specific analysis of ϕ torsions are nearly all within the experimental error of J-coupling determination. A more sensitive test would therefore refer to coupling data related to the side-chain torsion χ_1. However, NMR studies of proteins to date do not present sufficient a number of coupling values and types for self-consistent fitting to work satisfactorily. At least four coupling constants are required per residue, and a sufficiently large number of residues is needed also, such that there would be many more data available than parameters to optimize. Phase redundancy between various pairs of Karplus curves for χ_1, especially, complicates the analysis in addition, requiring additional data to resolve the ambiguities.

Sufficiently large homogeneous $^3J(\chi_1)$-coupling datasets not being available from proteins other than flavodoxin, Karplus coefficients presented in Table 5 were gauged, using various sample data reported in the literature, by applying a more conventional procedure which included the straightforward calculation and comparison of two values for the RMSD_J between sets of predicted and experimental coupling constants. Predicted coupling constants were obtained, on the one hand, using those amino-acid independent Karplus parameters employed in the respective protein study published and, on the other hand, using the amino-acid specific Karplus parameters presented here. In both cases, predictions were made on the basis of known χ_1 torsion angle parameters as obtained from well-resolved crystal coordinates. Without detailed inspection of the results, a smaller RMSD_J in the second case would indicate that the new Karplus coefficients were superior in translating between χ_1 torsion angle and ^3J coupling contants.

For example, an investigation into the structure of hen egg-white lysozyme provided experimental $^3J_{H\alpha,H\beta}$ coupling constants for 57 residues out of the 129-amino acid chain [29]. The same publication also quotes reference χ_1 torsion angles obtained from the tetragonal type-2 crystal structure of the enzyme resolved at 2.0 Å. The investigators employed Karplus parameters as given by DeMarco et al. [30] and calculated RMSD_J from only those 23 residues for which unambiguous stereospecific H^β shift assignments could be made (as given in their Table 4). Concentrating, for the purpose of testing a wider amino-acid type range, on the subset of those 41 residues that do not exhibit any apparent χ_1 angular dynamics, and for which stereospecific H^β shift assignments were either unambiguous (as given in their Table 4) or chosen best to agree with X-ray data (as discussed in their Fig. 4), the value of RMSD_J obtained is 1.59 Hz. Smith et al. [29] stress the importance of

Table 7
Amino-acid ϕ related Karplus coefficients self-consistently inferred from flavodoxin ^3J couplings[a]

J type	ϕ substituent pattern	C_0	C_1	C_2	A	B	C	J_{trans}	J_{gauche}
	Ala	5.20					1.15	10.25	2.68
	other amino acids	5.15					1.10	10.20	2.62
	Ile, Val	5.10					1.05	10.14	2.57
$^3J(H^N,H^\alpha)$	**Consensus**	**5.05**	**−1.00**	**4.05**	**8.09**	**−1.00**	**1.01**	**10.10**	**2.53**
	Ser	4.88					0.84	9.93	2.36
	Thr	4.83					0.78	9.88	2.31
	Cys	4.73					0.68	9.78	2.21
	Fundamental	*4.69*					*0.64*	*9.74*	*2.17*
	Gly	4.59					0.54	9.64	2.06
	Ala	2.11					−0.31	5.59	0.37
	other amino acids	2.08					−0.33	5.56	0.34
	Ile, Val	2.06					−0.36	5.53	0.32
$^3J(H^N,C')$	**Consensus**	**2.03**	**−1.06**	**2.42**	**4.83**	**−1.06**	**−0.38**	**5.51**	**0.30**
	Ser	1.95					−0.47	5.43	0.21
	Thr	1.92					−0.49	5.40	0.18
	Cys	1.87					−0.54	5.35	0.13
	Gly	1.80					−0.61	5.28	0.06
	Fundamental	*1.64*					*−0.78*	*5.12*	−0.10
	Ser	1.87					0.39	3.67	0.97
	Cys	1.85					0.36	3.64	0.95
	Fundamental	*1.79*					*0.31*	*3.59*	*0.89*
	Thr	1.76					0.28	3.56	0.86
	Ala	1.74					0.26	3.53	0.84
	Met	1.63					0.15	3.43	0.73
$^3J(H^N,C^\beta)$	**Consensus**	**1.59**	**−0.31**	**1.47**	**2.95**	**−0.31**	**0.11**	**3.39**	**0.69**
	Arg, Glx, Lys, Pro	1.57					0.08	3.36	0.67
	Val	1.52					0.04	3.32	0.62
	Leu	1.50					0.02	3.30	0.60
	His	1.47					−0.01	3.27	0.57
	Asn	1.46					−0.02	3.26	0.56
	Ile	1.46					−0.03	3.25	0.56
	Phe, Trp, Tyr	1.44					−0.04	3.24	0.56
	Asp	1.43					−0.05	3.23	0.54
	Pro	3.46					1.17	7.87	1.26
	Ala	3.18					0.89	7.59	0.98
	other amino acids	3.16					0.87	7.57	0.95
	Ile, Val	3.13					0.84	7.54	0.92
$^3J(C_{i-1}, H_i^\alpha)$	**Consensus**	**3.11**	**−2.12**	**2.29**	**4.58**	**−2.12**	**0.82**	**7.52**	**0.90**
	Ser	3.02					0.73	7.43	0.82
	Thr	3.00					0.71	7.41	0.79
	Cys	2.95					0.66	7.36	0.74
	Gly	2.88					0.59	7.29	0.67
	Fundamental	*2.83*					*0.54*	*7.24*	*0.62*
	Pro	1.30					0.44	3.24	0.33
	Ala	1.18					0.31	3.12	0.21
	other amino acids	1.16					0.30	3.10	0.19
	Ile, Val	1.15					0.29	3.09	0.18
$^3J(C_{i-1}, C_i')$	**Consensus**	**1.14**	**−1.08**	**0.86**	**1.73**	**−1.08**	**0.28**	**3.08**	**0.17**
	Ser	1.10					0.23	3.04	0.13
	Thr	1.08					0.22	3.02	0.11
	Cys	1.06					0.20	3.00	0.09
	Gly	1.02					0.16	2.96	0.05
	Fundamental	*0.89*					*0.03*	*2.83*	−0.08
	Pro	1.86					0.52	3.81	0.89
	Ser	1.86					0.51	3.81	0.89
	Cys	1.85					0.50	3.80	0.88
	Thr	1.81					0.46	3.75	0.83
	Ala	1.79					0.45	3.74	0.82
	Fundamental	*1.77*					*0.42*	*3.72*	*0.80*
	Met	1.74					0.39	3.69	0.77

Table 7 (*continued*)

J type	ϕ substituent pattern	C_0	C_1	C_2	A	B	C	J_{trans}	J_{gauche}
$^3J(C'_{i-1}, C_i^\beta)$	**Consensus**, Thr	**1.72**	**−0.60**	**1.35**	**2.69**	**−0.60**	**0.37**	**3.67**	**0.75**
	Arg, Glx, Lys	1.71					0.36	3.65	0.73
	Val	1.69					0.34	3.63	0.71
	Leu	1.68					0.33	3.62	0.70
	Asn	1.66					0.31	3.60	0.68
	Ile	1.65					0.31	3.60	0.68
	His, Phe, Trp, Tyr	1.64					0.30	3.59	0.67
	Asp	1.64					0.29	3.59	0.67

[a] *Notes:* Details apply as given in Table 5 footnote, except that $\theta = \phi + \Delta\phi$ and *consensus* coefficients are averages over 135 residues: Ala (17), Arg (7), Asn (2), Asp (16), Cys (4), Gln (3), Glu (12), Gly (13), His (1), Ile (9), Leu (12), Lys (4), Met (0), Phe (6), Pro (0), Ser (8), Thr (5), Trp (2), Tyr (5), Val (9).

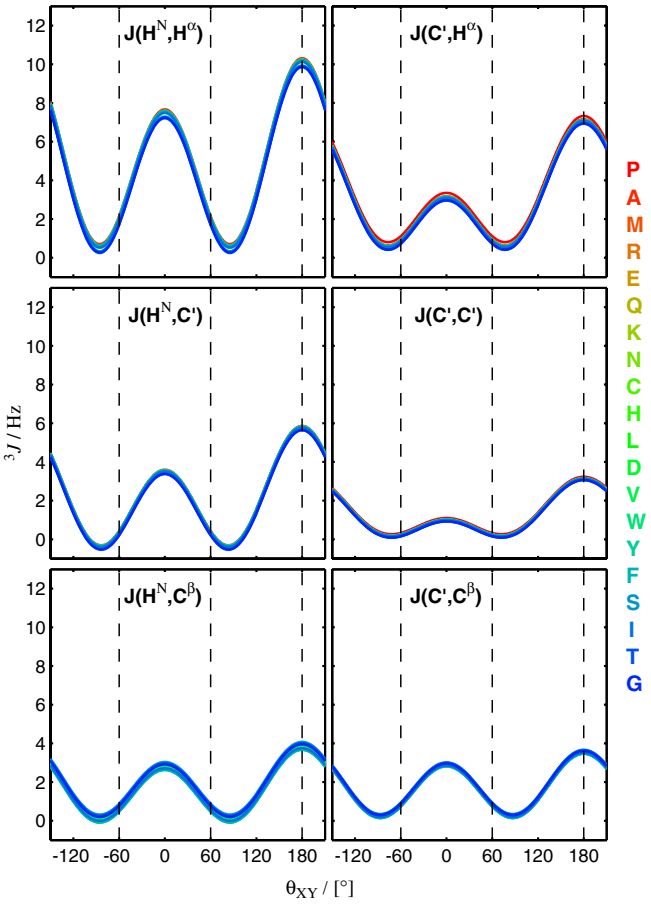

Fig. 4. Amino-acid type specific Karplus curves for all main-chain $J(\phi)$ couplings. Angle-independent offsets result from self-consistent modeling including Gaussian angular motion and second-sphere substituent effects (M_1S_2). It is seen that the spread due to amino-acid type barely exceeds the eperimental precision.

applying a correction factor according to Kopple et al. [31] in order to account for electronegativity of the side-chain oxygen atoms in serine and threonine, effectively reducing the respective calculated coupling values by 8% and improving the $RMSD_J$ to 1.52 Hz. Applying the new amino-acid specific Karplus parameters for $^3J_{H\alpha,H\beta}$ developed in the present work, which take into account the various side-chain topologies, was found further to improve the discrepancy in the lysozyme example to yield an $RMSD_J$ of 1.33 Hz.

Similarly, experimental $^3J_{H\alpha,H\beta}$ and $^3J_{N',H\beta}$ coupling constants available for the FKBP/FK506 complex [32] were subjected to an analysis of $RMSD_J$ between predicted and experimental coupling values. Referencing torsion angles (also given in the same paper) as obtained from the X-ray structure with PDB accession code 1FKF [33] and commonly used Karplus parameters for the χ_1 torsion-angle dependencies of $^3J_{H\alpha,H\beta}$ [30] and $^3J_{N',H\beta}$ [34] resulted in values of $RMSD_J$ of 2.89 (94 data) and 1.45 Hz (103 data), respectively. Both these discrepancies between predicted and experimental coupling constants decreased to 2.42 and 1.10 Hz, respectively, when the Karplus parameters developed in this work were applied.

5. Conclusions

Disentangling the subtle differential effects from relocating any given substituent provides the key to a more quantitative understanding of the topology dependency of J couplings, ultimately resulting in bespoke sets of Karplus coefficients applicable to the molecular fragment in question. The quantitative determination of substituent-related component couplings has been carried out self-consistently, on the exclusive basis of experimental data that is. The versatility of the proposed concept allows suitable Karplus equations to be constructed for other conceivable J coupling types, not necessarily only polypeptide related ones.

The two sets of component couplings derived here for χ_1 (Table 3) and for ϕ (Table 4) must not be compared directly because the different nature of the intervening atoms in the coupling paths causes fundamental and incremental effects to partition differently. It is worth emphasizing, though, that the entirety of Karplus coefficients obtained for each respective torsion type has wholly undergone the same self-consistent calibration procedure and so forms a homogeneous set with comparable confidence margins.

The component couplings presented here for polypeptides are tentative in so far as their iterative nature makes them subject to change in the experimental data pool. The variable nature of parameter optimization would make it necessary to optimize anew each respective protein dataset to see whether the parameters, i.e., the Karplus coefficients, obviously not the torsion angles, are similar in other cases. Expectations are that the parameters would

not come out dramatically differently with other proteins. The general tendencies of the substituent effects outlined are likely to prevail in future high-precision ^3J studies on other proteins.

Acknowledgment

Frank Löhr is thanked for stimulating discussion and his continual interest in this *J*-coupling study.

References

[1] P. Güntert, C. Mumenthaler, K. Wüthrich, Torsion angle dynamics for NMR structure calculation with the new program DYANA, J. Mol. Biol. 273 (1997) 283–298.
[2] M. Karplus, Vicinal proton coupling in nuclear magnetic resonance, J. Am. Chem. Soc. 85 (1963) 2870–2871.
[3] M. Karplus, The analysis of molecular wave functions by nuclear magnetic resonance spectroscopy, J. Phys. Chem. 64 (1960) 1793–1798.
[4] V.F. Bystrov, Spin–spin coupling and the conformational states of peptide systems, Prog. NMR Spectr. 10 (1976) 41–81.
[5] C.A.G. Haasnoot, F.A.A.M. de Leeuw, C. Altona, The Relationship between proton–proton NMR coupling constants and substituent electronegativities- I. An empirical generalization of the Karplus equation, Tetrahedron 36 (1980) 2783–2792.
[6] C.A.G. Haasnoot, F.A.A.M. de Leeuw, H.P.M. de Leeuw, C. Altona, Relationship between proton–proton NMR coupling constants and substituent electronegativities. III. Conformational analysis of proline rings in solution using a generalized Karplus equation, Biopolymers 20 (1981) 1211–1245.
[7] J.M. Schmidt, M. Blümel, F. Löhr, H. Rüterjans, Self-consistent ^3J coupling analysis for the joint calibration of Karplus coefficients and evaluation of torsion angles, J. Biomol. NMR 14 (1999) 1–12.
[8] C. Pérez, F. Löhr, H. Rüterjans, J.M. Schmidt, Self-consistent Karplus parametrization of ^3J couplings depending on the polypeptide sidechain torsion χ_1, J. Am. Chem. Soc. 123 (2001) 7081–7093.
[9] M.A. Knauf, F. Löhr, G.P. Curley, P. O'Farrell, S.G. Mayhew, F. Müller, H. Rüterjans, Homonuclear and heteronuclear NMR studies of oxidized *Desulfovibrio vulgaris* flavodoxin: sequential assignments and identification of secondary structure, Eur. J. Biochem. 213 (1993) 167–184.
[10] M.A. Knauf, F. Löhr, M. Blümel, S.G. Mayhews, H. Rüterjans, NMR investigation of the solution conformation of oxidized flavodoxin from *Desulfovibrio vulgaris*: determination of the tertiary structure and detection of protein-bound water molecules, Eur. J. Biochem. 238 (1996) 423–434.
[11] W. Watt, A. Tulinsky, R.P. Swenson, K.D. Watenpaugh, Comparison of the crystal structures of a flavodoxin in its three oxidation states at cryogenic temperatures, J. Mol. Biol. 218 (1991) 195–208.
[12] M.A. Walsh, Ph.D. Thesis, National University of Ireland, 1994.
[13] M.A. Walsh, A. McCarthy, P.A. O'Farrell, P. McArdle, P.D. Cunningham, S.G. Mayhew, T.M. Higgins, X-ray crystal structure of the *Desulfovibrio vulgaris* (Hildenborough) apoflavodoxin–riboflavin complex, Eur. J. Biochem. 258 (1998) 362–371.
[14] O. Jardetzky, On the nature of molecular conformations inferred from high-resolution NMR, Biochem. Biophys. Acta 621 (1980) 227–232.
[15] J.C. Hoch, C.M. Dobson, M. Karplus, Vicinal coupling constants and protein dynamics, Biochemistry 24 (1985) 3831–3841.
[16] Y. Karimi-Nejad, J.M. Schmidt, H. Rüterjans, H. Schwalbe, C. Griesinger, Conformation of valine side chains in ribonuclease T_1 determined by NMR studies of homonuclear and heteronuclear ^3J coupling constants, Biochemistry 33 (1994) 5481–5492.
[17] R. Brüschweiler, D.A. Case, Adding harmonic motion to the Karplus Relation for spin–spin coupling, J. Am. Chem. Soc. 116 (1994) 11199–11200.
[18] J.M. Schmidt, Conformational equilibria in polypeptides. II. Dihedral-angle distribution in antamanide based on three-bond coupling information, J. Magn. Reson. 124 (1997) 310–322.
[19] L.B. Krivdin, E.W. Della, Spin–spin coupling constants between carbons separated by more than one bond, Prog. NMR Spectr. 23 (1991) 301–610.
[20] P.E. Hansen, Long range ^{13}C–^{13}C coupling constants. A review, Org. Magn. Reson. 5 (1978) 215–233.
[21] P.E. Hansen, Carbon–hydrogen spin–spin coupling constants, Prog. NMR Spectr. 14 (1981) 175–296.
[22] J.A. Pople, J.W. McIver Jr., N.S. Ostlund, Self-consistent perturbation theory. II. Nuclear-spin coupling constants, J. Chem. Phys. 49 (1968) 2965–2970.
[23] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes, Cambridge University Press, Cambridge, 1989.
[24] M.L. Severson, G.E. Maciel, A molecular orbital study of the dihedral angle dependencies of vicinal carbon–carbon coupling constants, J. Magn. Reson. 57 (1984) 248–268.
[25] M. Blümel, J.M. Schmidt, F. Löhr, H. Rüterjans, Quantitative ϕ torsion angle analysis in *Desulfovibrio vulgaris* Flavodoxin based on six ϕ related ^3J couplings, Eur. Biophys. J. 27 (1998) 321–334.
[26] M. Barfield, J.L. Marshall, E.D. Canada, Nuclear spin–spin coupling via nonbonded interactions. 2. γ-Substituent effects for vicinal coupling constants involving ^{13}C, J. Am. Chem. Soc. 102 (1980) 7–12.
[27] F. Löhr, J.M. Schmidt, M. Maurer, H. Rüterjans, Improved measurement of $^3J(H_i^\alpha, N_{i+1})$ coupling constants in H_2O dissolved proteins, J. Magn. Reson. 153 (2001) 75–81.
[28] L.A. Donders, F.A.A.M. de Leeuw, C. Altona, Relationship between proton–proton NMR coupling constants and substituent electronegativities, Magn. Res. Chem. 27 (1989) 556–563.
[29] L.J. Smith, M.J. Sutcliffe, C. Redfield, C.M. Dobson, Analysis of ϕ and χ_1 torsion angles for Hen Lysozyme in solution from ^1H NMR spin–spin coupling constants, Biochemistry 30 (1991) 986–996.
[30] A. DeMarco, M. Llinás, K. Wüthrich, Analysis of the ^1H-NMR spectra of Ferrichrome peptides. I. The non-amide protons, Biopolymers 17 (1978) 617–636.
[31] K.D. Kopple, G.R. Wiley, R. Tauke, A dihedral angle-vicinal proton coupling constant correlation for the $\alpha-\beta$ bond of amino acid residues, Biopolymers 12 (1973) 627–636.
[32] R.X. Xu, E.T. Olejniczak, S.W. Fesik, Stereospecific assignments and χ_1 rotamers for FKBP when bound to ascomycin from $^3J_{H\alpha,H\beta}$ and $^3J_{N,H\beta}$ coupling constants, FEBS Lett. 305 (1992) 137–143.
[33] G.D. Van Duyne, R.F. Standaert, P.A. Karplus, S.L. Schreiber, J. Clardy, Atomic structure of FKBP-FK506, an immunophilin-immunosuppressant complex, Science 252 (1991) 839–842.
[34] A. DeMarco, M. Llinás, K. Wüthrich, ^1H–^{15}N spin–spin couplings in alumichrome, Biopolymers 17 (1978) 2727–2742.