Received: 20 April 2022 | Revised: 17 June 2022 | Accepted: 28 June 2022

DOI: 10.1002/qua.26979

RESEARCH ARTICLE

International Journal of Quantum Chemistry WILEY

Searching conformational analysis of Asp residues through theoretical ^3J vicinal coupling constants and Karplus equations

María Botella | Ignacio Ema | Jesús San Fabián | Jose Manuel García de la Vega

Departamento de Química Física Aplicada, Universidad Autónoma de Madrid, Facultad de Ciencias, Madrid, Spain

Correspondence
Jesús San Fabián, Departamento de Química Física Aplicada, Universidad Autónoma de Madrid, Facultad de Ciencias, 28049, Madrid, Spain.
Email: jesus.sanfabian@uam.es

Abstract

The dependence of the vicinal spin–spin coupling constants on the torsion side chain angle χ_1 through the Karplus equations is considered for the study of the structure of Asp amino acid residues. Experimental and theoretical, obtained with density functional theory methods, vicinal coupling constants combined with extended Karplus equations, which include six coefficients, are applied to a dipeptide model of the amino acid Asp. To find out the empirical χ_1 angles of the side chain, a statistical analysis procedure is developed to compute the rmsd values and find the χ_1 as the minimum of those values. The χ_1 values obtained in this work for nine Asp residues of the flavodoxin protein *Desulfovibrio vulgaris* are successfully compared with those derived by nuclear magnetic resonance and X-rays.

KEYWORDS
Asp dipeptide model, density functional theory, Karplus equations, side-chain conformation, spin–spin coupling constants

1 | INTRODUCTION

Knowledge of protein structures is essential to understand their activity and biochemical processes [1]. Proteins have thousands of atoms and complex shapes. The direct relationship between the structure of the protein and its functions or properties makes the study of geometry an important topic. Information about the conformations of proteins and their amino acid sequence helps to better understand the functions of proteins [2]. The knowledge of the side chain dihedral angles, from relaxation experiments, provides us with information on the structure of proteins and, consequently, on their dynamics [3]. These angles adopt generally staggered structures, corresponding to certain values, while some protein residues prefer certain combination of them [4, 5].

The determination of molecular structures of proteins is performed with different characterization techniques in gas, liquid and solid phases: in gas phase, microwave spectroscopy or electron diffraction; in liquid phase, nuclear magnetic resonance (NMR); and in solid phase, X-ray diffraction. NMR is the only method that allows the determination of the 3D structure and dynamics of proteins in solution (physiological medium) [6]. Information obtained from NMR spectra includes chemical shifts, coupling constants, and nuclear Overhauser effects. All these factors are useful for the analysis of conformational properties [7–9].

Asp is a non-essential amino acid that participates in protein synthesis acting as a neurotransmitter and can be found in its D and L isomers [1]. It belongs to the group of AAs with polar side chains since it has a terminal carboxyl group that can appear as ionized. This carboxyl group can release a proton and, hence, acquire a negative charge at the pH of body fluids. Proteins and peptides are degraded, and their biological activity is changed by many chemical reactions, such as deamidation, hydrolysis, oxidation, and racemization, as well as, other factors like pH. Asp residues undergo deamidation via the formation of succinimide intermediates at neutral and slightly alkaline pH, resulting in the formation of isoaspartate

(β-AA and aspartic acid residue isomeric) linkages and racemized aspartate residues [10]. Moreover, it plays an important role in the synthesis of other AAs and in the citric acid and urea cycles.

Flavodoxins are electron transfer proteins involved in a variety of photosynthetic reactions in bacteria and contain one molecule of non-covalently but tightly bound flavin mononucleotide as the redox active component [11]. The biophysical interest in protein stability, folding and ligand recognition has stimulated the use of several flavodoxins as models to understand the specific details of the pathway by which the synthesized flavodoxin polypeptide becomes functional and also the fundamentals of these processes, which are common to every protein [12, 13]. The combination of experimental methods, such as NMR, and theoretical methods based on the density functional theory (DFT) allows to estimate the value of the side chain angle of the AAs in solution, obtaining results similar to those produced by experimental analysis [14]. Vicinal spin–spin coupling constants, (SSCCs) $^3J_{XY}$ are scalar couplings that assume an interaction between the spins of the XY nuclei separated by three bonds (X—C—C—Y). These coupling constants depend on the χ_1 torsion angle of the selected AA [15–18], in this case the Asp dipeptide. To obtain theoretically Karplus equation, that is, the Fourier coefficients, is essential to calculate a selection of optimized geometries of Asp dipeptide. Once the Fourier coefficients are theoretically determined, we can predict the coupling constants of the side chain in the protein and obtain the torsion angles χ_1 which are compared with those derived experimentally. The final objective of the present work is to present an improved procedure for determining side chain torsional angles using a combination of theoretical and experimental results and to apply it for determining the molecular structure of the side chain of the aspartic (Asp) AA in the Flavodoxin of *Desulfovibrio vulgaris* (strain Hildenborough).

2 | THEORY AND MODELS

2.1 | Karplus equation

The Karplus equation [19, 20] relates the vicinal coupling constants with the dihedral or torsion angles between the X—C—C—Y coupling atoms, that is, the angle (θ) formed by the X—C—C and C—C—Y planes allowing us to analyze their rotation, for example, in the AAs of proteins. This dependency was established under the following expression:

$$^3J_{XY} = C_0 + C_1\cos(\theta) + C_2\cos(2\theta) \tag{1}$$

The Karplus equation, Equation (1), was originally restricted to three Fourier coefficients (C_0, C_1, C_2) due to its empirical limitations and was theoretically derived from calculations for the HCCH fragment [19, 20]. From the comparison between the theoretical and experimental data, the dependence of $^3J_{XY}$ on other additional molecular parameters such as the electronegativity and the relative position of the substituents and the local geometry (bond lengths and bond angles) was also observed and indicated long ago [20]. Consequently, we consider an extended Fourier development with more terms and for specific fragments that includes the relationship between the vicinal coupling constants and the dihedral angle of the side chain χ_1, together with the substituent effects. In this way, errors are, at least partially, minimized.

$$^3J_{XY} = C_0 + \sum_{n=1}^{P} C_n\cos(n\theta) + \sum_{n=1}^{Q} S_n\sin(n\theta) \tag{2}$$

Theoretical and empirical studies [21–23] show that in Equation (2), the necessary values of P and Q can be limited to $P = 3$ and $Q = 2$, generating a total of six Fourier coefficients (C_0, C_1, C_2, C_3, S_1, S_2).

$$\begin{aligned} ^3J_{XY}^{cal} \;=\; & C_0 + C_1\cos(\theta) + C_2\cos(2\theta) + C_3\cos(3\theta) \\ & + S_1\sin(\theta) + S_2\sin(2\theta) \end{aligned} \tag{3}$$

This series of six coefficients determined for a specific AA is considered to include the effect of the nature and orientation of the substituents linked to the side chain and the effects of local geometry. Once the Fourier coefficients have been obtained for each AA model and in an inverse process, they can be used to predict the conformation of the side chain geometry [17, 18, 23].

2.2 | Selection of the model

For studying the geometry of the side chain and the SSCCs in Asp residues, it is necessary to select a molecular model that must incorporate most of the studied properties and that in turn should be valid, considering the computational resources, for the chosen computational protocol. The combination of AAs generates peptides and from them proteins that arrange in the space mainly in two different secondary structures: α-helix or

β-sheet. These two conformations differ in the angles of the planes that are formed around the C^{α} (backbone) with $\varphi \approx -64^{\circ}$ and $\psi \approx -44^{\circ}$ for α-helix and $\varphi \approx -121^{\circ}$ and $\psi \approx \pm 128^{\circ}$ for β-sheet (Figure 1). Furthermore, these arrangements are generated by steric and/or non-covalent interactions between the AA residues of the protein; therefore, it is necessary to select a simple and adequate model that does not influence the final objective of this work. The limitation in the size of the model for carrying out the theoretical calculations makes us lose non-covalent interactions (e.g., hydrogen bonds), which would cause the secondary structure of the protein. We set the angles φ and ψ during the geometry optimization for maintaining the secondary structure and to make our results applicable to the protein.

The chosen dipeptide, shown in Figure 1, represents a simplified model [24] that incorporates the effect of substituents on the coupling constants. In order to produce a successful study, molecular parameters of the model are optimized fixing or restricting some of them. Clearly, the optimization of this dipeptide model generates a geometry that is quite far from that of the secondary structure of proteins, α-helix and β-sheet. In a previous work on Ala, the differences between the α-helix and β-sheet conformations in the SSCCs were estimated at a rmsd of 15% [23]. Due to the greater presence in the residues and the similarity they present with the experimental values, the α-helix is selected as the most appropriate. Accordingly, we decided to set the characteristic angles of α-helix [25, 26] for the dipeptide model with the values of the angles $\varphi = -64$ and $\psi = -44^{\circ}$. The dipeptide model does not account for long-range effects or the interactions of other residues. Therefore, a full optimization of this model would generate a conformation different from that usually presented in proteins.

2.3 | SSCCs and Fourier coefficients

Once the model is selected and the geometries are generated, the vicinal SSCCs around the C^{α}—C^{β} bond (Figure 1) are calculated for the Asp with 6 of the 12 optimized geometries. As shown previously [23], it is not necessary to use all 12 geometries considering the small differences obtained and that the computational effort is large. These calculations generate nine SSCCs for each optimized geometry: ${}^{3}J_{H^{\alpha}H^{\beta_2}}$, ${}^{3}J_{H^{\alpha}H^{\beta_3}}$, ${}^{3}J_{H^{\alpha}C'}$, ${}^{3}J_{C'H^{\beta_2}}$, ${}^{3}J_{C'H^{\beta_3}}$, ${}^{3}J_{C'C'}$, ${}^{3}J_{N'H^{\beta_2}}$, ${}^{3}J_{N'H^{\beta_3}}$, and ${}^{3}J_{N'C'}$ (see Newman projection in Figure 1). With six values for each constant, a group of six coefficients, C_0, C_1, C_2, C_3, S_1, and S_2, are obtained. These Fourier coefficients were derived by interpolation to Equation (3) using the calculated J couplings. An interpolation was chosen instead of an inverse Fourier analysis because the dihedral angles are not exactly evenly spaced. In a Fourier analysis with equally spaced values it is not necessary to include more than six values to make the fit since the higher order coefficients are usually small and do not affect the final results [23, 27]. Thus, nine sets of Fourier coefficients were obtained for the nine types of constants.

For each residue, a statistical parameter rmsd is obtained by comparing the experimental SSCCs [28] with those calculated using the Karplus equation. The angle χ_1 is then estimated by finding the minimum of the rmsd versus χ_1. Biochemical statistics define this parameter as the structural difference of two topologies. In this way, for a lower value, the similarity between the two structures is greater. We define rmsd as:

$$\text{rmsd}\ (\chi_1) = \sqrt{\frac{\sum_{i=1}^{n}\left(J_i^{\text{exp}} - J_i^{\text{cal}}(\chi_1)\right)^2}{n}} \tag{4}$$

where n is the number of experimental SSCCs, J_i^{exp}, and J_i^{cal} are those calculated using Equation (3). This equation is applied individually for each residue studied.

The above equation assumes the existence of a dominant conformer for each residue. Another possibility is to consider the existence of three conformers, corresponding to alternating geometries. Although this approach does not allow to obtain the precise value of the

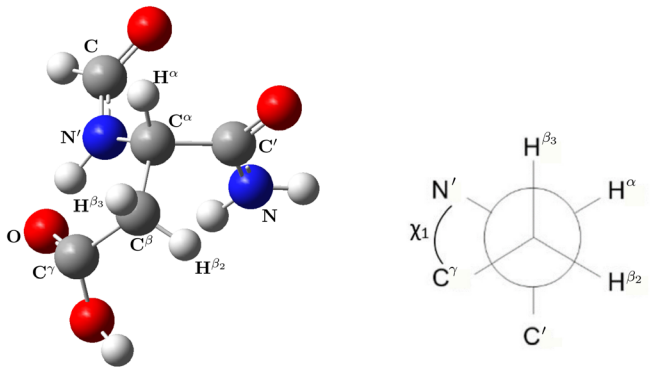

FIGURE 1 Molecular Asp dipeptide model. Angle definitions: φ (C—N′—C^{α}—C′), ψ (N′—C^{α}—C′—N), χ_1 (N′—C^{α}—C^{β}—C^{γ}) and χ_2 (C^{α}—C^{β}—C^{γ}—O). Newman's projection indicating the dihedral angle χ_1

angle χ_1, it enables to analyze the dynamics of the system as the existence of several conformers. In this second approach the rmsd is define as:

$$\text{rmsd} = \sqrt{\frac{\sum_{i=1}^{n}\left[J_i^{\text{exp}} - \left(P_{60}\, J_i^{\text{cal}}(60) + P_{180}\, J_i^{\text{cal}}(180) + P_{300}\, J_i^{\text{cal}}(300)\right)\right]^2}{n}} \tag{5}$$

where P_X and $J_i^{\text{cal}}(X)$ are, respectively, the population of conformers, and the vicinal coupling calculated for the indicated angles (X = 60°, 180°, and 300°). Using the experimental and the calculated SSCCs, the population for the three staggered conformer can be obtained finding the minimum in the rmsd define in the above equation.

2.4 | Computational details

The geometry optimization process have been carried out at the B3LYP/6-31G(d,p) level [29–33] with Gaussian 09 program [34] and SSCCs calculations were performed at the B972/6-31++G**-J level [35, 36]. Both approaches for geometries and SSCCs have been used satisfactorily in previous studies [15, 23, 36]. All the results of SSCCs shown in this work are referred to isotopes ^{1}H, ^{13}C, and ^{15}N. As already mentioned, the selection of the fragment and the size of the model is very important because side chain calculated SSCCs must show high similarity to those of a real protein. The structure represented in Figure 1 is considered adequate for the calculation [23]. The conformation of the backbone is another factor that would affect the Fourier coefficients.

Both SSCCs and dihedral angles χ_1 are used to calculate the Fourier coefficients which include the molecular parameters that influence the coupling constants, such as the effect of the substituents (type, number and position) attached.

The work plan has been developed according to a computational protocol [23] summarized in the following steps:

1. Geometries were partially optimized by rotating the angle χ_1 in 30° intervals obtaining 12 optimized geometries for α-helix conformation ($\varphi = -64$ and $\psi = -44°$, respectively).
2. From the above geometries, six were selected (alternate and eclipsed conformations) and used to calculate, for each of them, nine SSCCs: $^3J_{H^\alpha H^{\beta_2}}$, $^3J_{H^\alpha H^{\beta_3}}$, $^3J_{H^\alpha C^\gamma}$, $^3J_{C'H^{\beta_2}}$, $^3J_{C'H^{\beta_3}}$, $^3J_{C'C^\gamma}$ $^3J_{N'H^{\beta_2}}$, $^3J_{N'H^{\beta_3}}$, and $^3J_{N'C^\gamma}$, involved around χ_1 (Figure 1).
3. From the previous SSCCs, six Fourier coefficients were calculated for each type of SSCC according to Equation (3) by interpolation.
4. Plotting rmsd (Equation 4) versus χ_1 (see below) allows to find the χ_1 value, that corresponding to the minimum.
5. The geometries determined in point 4 are compared with those empirically derived by other NMR or X-ray experiments.

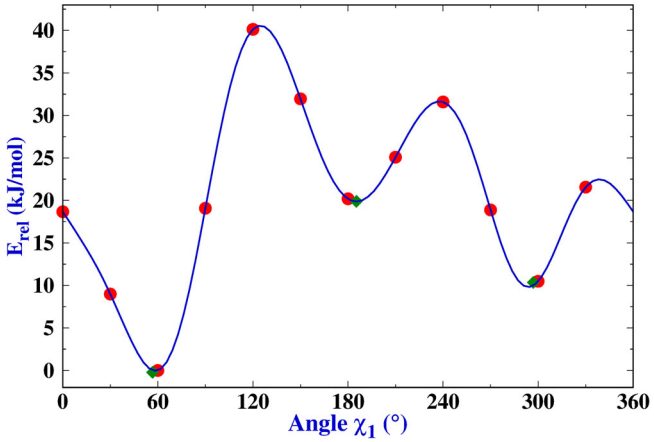

FIGURE 2 Relative energies (kJ/mol) versus the torsion angle of the side chain χ_1 from 0° to 360°

3 | RESULTS AND DISCUSSION

The side chain χ_1 angle (Figure 1) is rotated at 30° intervals covering the full rotation, while the α-helix conformation is fixed. Thus, a series of 12 conformations of optimized geometries has been obtained and their corresponding relative energies, with respect to χ_1, are represented in Figure 2. The most stable and unstable conformations of the dipeptide correspond to χ_1 torsion angles of ~60° and ~120°, respectively. Among the alternate conformations, the most stable one is that of ~60°, which is used as reference, followed by that of ~300° (around 10 kJ/mol higher) and that of ~180° (around 20 kJ/mol higher). The optimization of these three minima, that is, unconstraining the χ_1 angle, yield the angles, respectively, of 57°, 175°, and 297° which have been shown in Figure 2 (blue diamond). It should be noted that this order is not that empirically predicted using the SSCCs (see below).

To calculate and study the vicinal SSCCs, we take the eclipsed ($\chi_1 = 0°$, 120°, and 240°) and alternate ($\chi_1 = 60°$, 180°, and 300°) conformations. These six chosen conformations are shown in Figure 3. The nine computed vicinal coupling constants for the six chosen conformations are presented in Table 1. Vicinal SSCCs depend on the torsion angle χ_1 of the side chain and on the substituents (type and position) attached to the central atoms [27].

SSCCs are plotted versus the dihedral angle θ between the coupled nuclei in Figure 4 and fitted to a Fourier series with six different terms using Equation (3). For each group of SSCCs, six Fourier coefficients are obtained (see Table 2). These six Fourier coefficients (C_0, C_1, C_2, C_3, S_1, and S_2) generate significant differences in the values of the SSCCs with respect to those calculated with the classical Karplus equation with only three coefficients (C_0, C_1, C_2). In this regard, it is important to note that sometimes the value of S_2 is greater than that of C_1.

For comparison, Table S1 shows the values of six sets of empirical Fourier coefficients derived by Pérez et al. [28] using the three coefficients C_0, C_1, and C_2 of the classical Karplus equation, Equation (1). These empirical Fourier coefficients show similarity with our calculated values (see Tables 2 and S1). When making this comparison, it is important to note that Perez's empirical values only present coefficients for six SSCCs instead of nine as in our calculations, because they consider the H^{β_2} and H^{β_3} as equivalent.

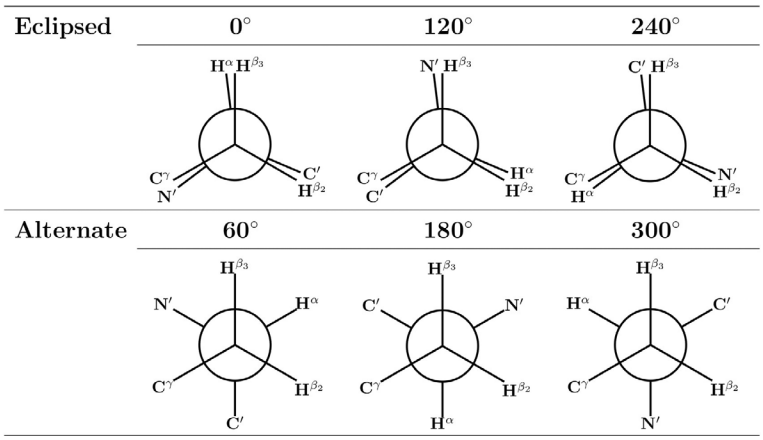

FIGURE 3 Newman's projections and χ_1 angle for the six selected conformations

TABLE 1 Vicinal SSCCs calculated (in Hz) for the six selected χ_1 (°) conformations

SSCC	0	60	120	180	240	300
$^3J_{H^\alpha H^{\beta_2}}$	3.30	3.17	12.9	5.11	2.69	12.6
$^3J_{H^\alpha H^{\beta_3}}$	11.6	5.85	1.62	12.3	3.93	3.31
$^3J_{H^\alpha C'}$	3.11	12.9	4.19	1.62	7.69	2.50
$^3J_{C'H^{\beta_2}}$	6.19	1.02	2.90	8.73	1.53	3.02
$^3J_{C'H^{\beta_3}}$	2.22	8.75	1.88	2.38	6.02	1.43
$^3J_{C'C'}$	1.60	1.10	3.00	0.60	2.00	6.30
$^3J_{N'H^{\beta_2}}$	−1.07	5.08	−1.79	−1.62	−5.46	−1.21
$^3J_{N'H^{\beta_3}}$	−1.25	1.17	−5.33	−1.28	−1.69	−5.49
$^3J_{N'C'}$	−0.29	0.23	−1.12	−3.70	−1.10	−0.20

Abbreviation: SSCC, spin–spin coupling constant.

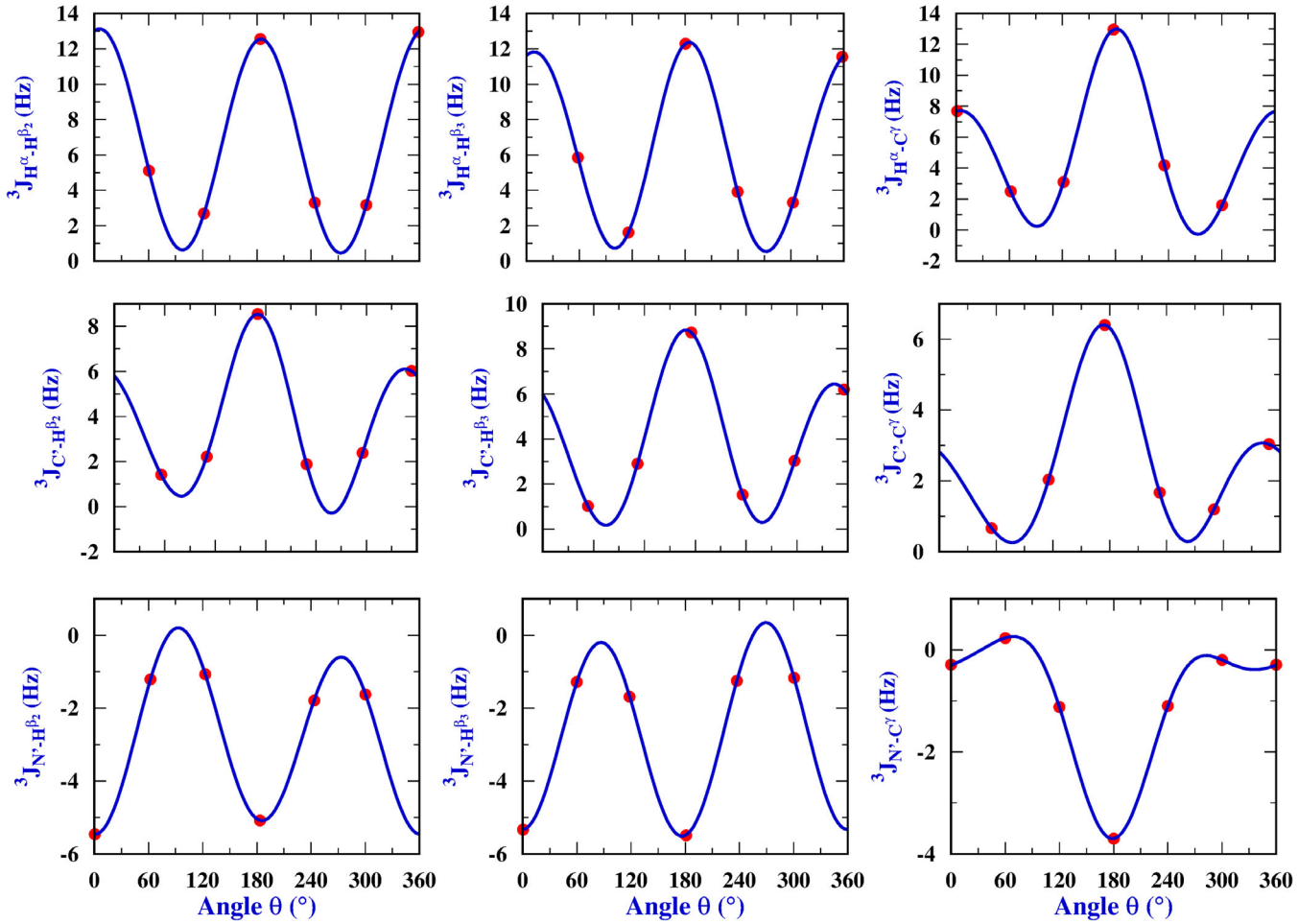

FIGURE 4 Extended Karplus equation representation as function θ angle

TABLE 2 Fourier coefficients (in Hz) using extended Karplus equation

SSCCs	C_0	C_1	C_2	C_3	S_1	S_2
$^3J_{H^\alpha H^{\beta 2}}$	6.70	0.46	6.04	−0.19	0.18	1.11
$^3J_{H^\alpha H^{\beta 3}}$	6.39	0.19	5.56	−0.52	0.28	1.24
$^3J_{H^\alpha C^\gamma}$	5.17	−2.08	5.17	−0.58	0.24	0.42
$^3J_{C'H^{\beta 2}}$	3.93	−1.05	3.31	−0.21	0.04	−1.64
$^3J_{C'H^{\beta 3}}$	3.70	−0.83	3.35	−0.41	0.32	−1.36
$^3J_{C'C^\gamma}$	2.49	−1.38	2.08	−0.37	0.04	−0.75
$^3J_{N'H^{\beta 2}}$	−2.73	−0.16	−2.52	−0.05	0.40	−0.28
$^3J_{N'H^{\beta 3}}$	−2.67	0.12	−2.74	−0.04	−0.28	0.19
$^3J_{N'C^\gamma}$	−1.03	1.51	−0.97	0.19	0.12	0.13

Abbreviation: SSCCs, spin–spin coupling constants.

In protein studies, the dihedral angle θ, corresponding to the dihedral angle between the coupled nuclei, does not always coincide with χ_1. This angle is related to χ_1 such that $\theta = \chi_1 + \Delta\theta$, where $\Delta\theta$ has a value of 0°, 120°, or −120° due to tetrahedral symmetry. The corresponding $\Delta\theta$ values are shown in Table S2. These angles should be considered when calculating SSCCs with Karplus equations before comparing them with experimental values. $\Delta\theta$ and χ_1 values are used to obtain the θ angles that are introduced in the Karplus equations, with the calculated Fourier coefficients of Table 2, to obtain theoretical SSCCs values. These theoretical vicinal coupling constants are generated for all values of χ_1, from 0°

to 360° at intervals of 1°. Comparing these theoretical ${}^3J_i^{cal}(\chi_1)$ values with those, ${}^3J_i^{exp}$, obtained experimentally by Perez et al. [28], Table S3, the rmsd is calculated as a function of χ_1, Equation (4), and represented in Figure 5.

It can be seen that all the rmsd representations show two minima, the first one close to an alternate conformer (60°, 180°, or 300°) and the second one displaced 180° with respect to the first one. This second minimum corresponds to an eclipsed form (Figure 3), so we consider that the valid value would belong to an alternate conformation. These two minima are difficult to avoid and are due to the non-independence of the Karplus equations [15].

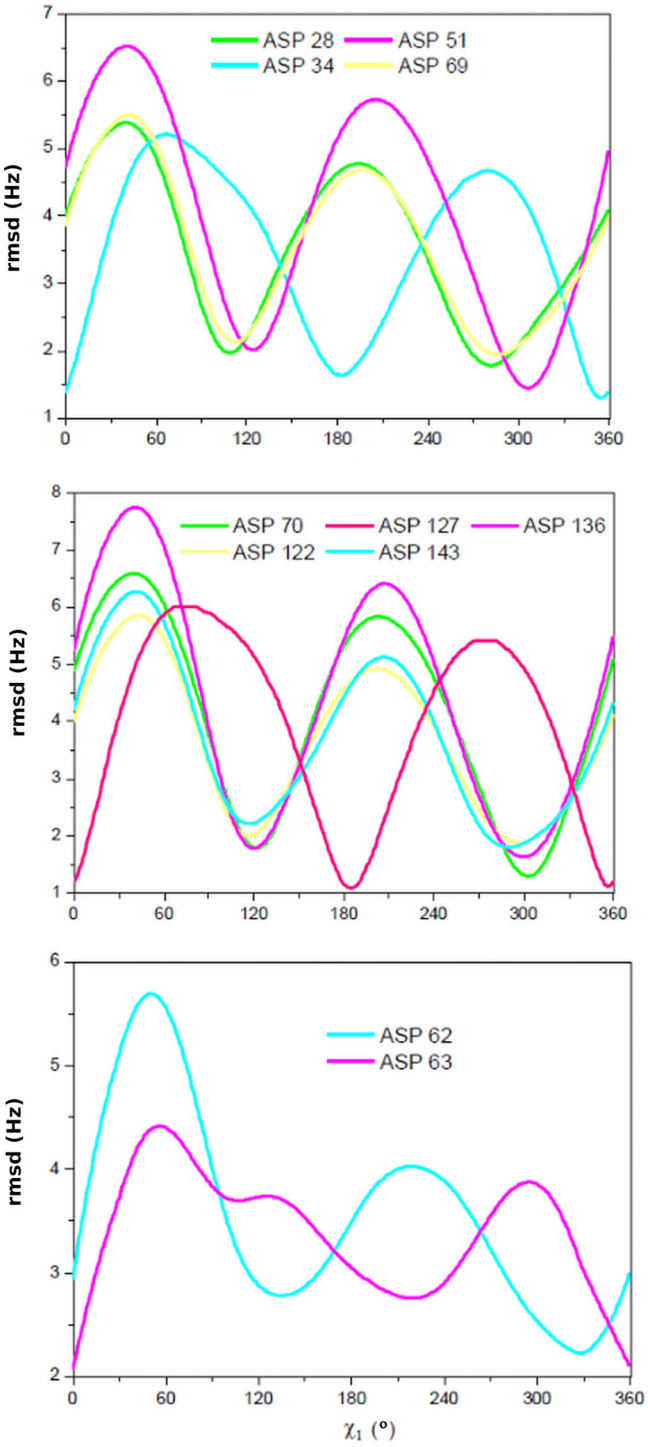

FIGURE 5 rmsd (Hz) versus χ_1 (degree) for some representative Asp residues

TABLE 3 Values of χ_1 (degree) of each minimum for Asp residues

Residue	Eclipsed angle	Alternate angle
Asp28	108	281
Asp34	355	184
Asp51	124	307
Asp62*	136	331
Asp63*	220	–
Asp69	111	288
Asp70	122	303
Asp122	116	296
Asp127	357	185
Asp136	119	300
Asp143	115	288

*Residues that do not show a defined conformation (see text).

TABLE 4 Populations (%) for conformations (60°, 180°, and 300°) of Asp residues

	Conformer		
Residue	60°	180°	300°
Asp28	24	9	**66**
Asp34	13	**71**	16
Asp51	4	15	**81**
Asp62	**13**	**34**	**53**
Asp63	**28**	**43**	**29**
Asp69	20	14	**66**
Asp70	5	11	**84**
Asp122	15	15	**70**
Asp127	10	**85**	5
Asp136	10	14	**77**
Asp143	14	16	**70**

Note: Bold values correspond to dominant population (larger than 60%) except for Asp62 and Asp63 where there is not a dominat one (see text).

A summary of the final results about χ_1 for each residue is shown in Table 3. There are seven residues with a χ_1 angle around 300° and two residues (Asp28 and Asp127) with an angle of 180°. Two other residues (Asp62 and Asp63) can be considered atypical because they do not show a defined conformation. This is because there is no single conformation that predominates, that is, two or three rotamers present non negligible populations. Therefore, these two last values have been represented separately in Figure 5 and will be omitted in the statistical comparison with other values previously obtained [28]. We find that it is necessary to carry out a population study in order to draw a clear conclusion on both residues. We have calculated the populations of three conformers 60°, 180°, and 300° for all residues using Equation (5). The comparison between the populations of 11 Asp residues is shown in Table 4. There are nine residues where a predominant population (larger than 66%) is clearly observed for one of the three conformers (angles 60°, 180°, or 300°). However for Asp62 and Asp63 residues is not possible to discern any dominant conformer for these three angles.

Data of side-chain χ_1 angles obtained in previous studies of NMR [28] and X-ray (PDB entries shown in Table S5) are compared with our predictions in Table S4. Three different values of rmsd encompassing different comparisons are also depicted in this Table S4, showing an excellent correlation between our results and experimental data [28]. The deviations of our theoretical χ_1 values and those empirically obtained are represented in Figure 6 for each Asp residues using the average X-ray values (Table S5) as a reference. The analysis of this Figure 6 and Table S4, where rmsd values versus those obtained as average of the X-ray results are also included, clearly shows the improvement of our results in seven residues. On the other hand, it is estimated that the side chain torsion angle χ_1 has a value close to 300° in most cases (seven out of nine) and in a small fraction (two out of nine, Asp34 and Asp127) it adopts an angle around 180°.

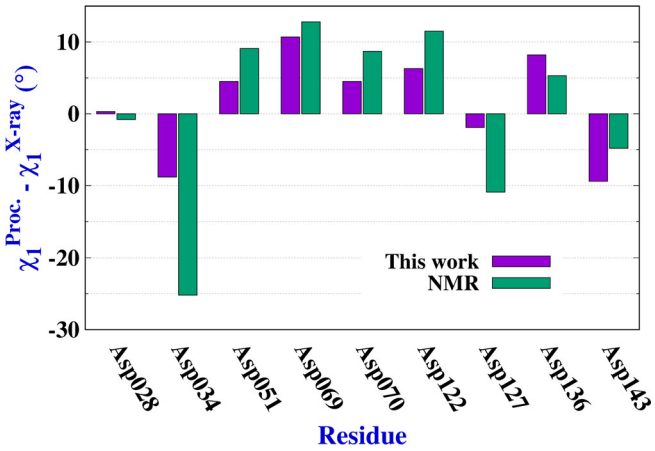

FIGURE 6 Deviations of χ_1 angles for Asp residues calculated with our procedure (this work) and nuclear magnetic resonance (NMR) [28] with respect to the average X-ray values

It is important to point out that the method applied in this work was able to identify highly mobile residues. Here the so-called atypical residues (Asp62 and Asp63) seem to present two or three conformers, respectively. The first one around 300° (53%) and 180° (34%) and then second one with populations of 28%, 43%, and 29% for 60°, 180°, and 300°, respectively.

4 | CONCLUSIONS

The χ_1 torsion angles of the side chain of Asp residues in the *D. vulgaris* flavodoxin have been estimated using a computational procedure. Asp dipeptide geometries have been optimized at the DFT level (B3LYP/6-31G(d,p)). Nine coupling constants for each selected optimized geometry were obtained at the B972/6-31++G**-J level, due to their favorable quality/cost ratio. The SSCCs were related to χ_1 using the Karplus equations that are based on a Fourier development with six coefficients. Analysis of theoretically optimized energies and geometries results in a χ_1 value of 60°. When this angle is predicted using theoretical Karplus equations and experimental SSCCs, the values are 180° or 300°, depending on the residue. This discrepancy may be due to the lack of non-bonding or long-range interactions in the theoretical dipeptide model when the optimization process is carried out. However, these interactions are implicitly included when the experimental SSCCs are used.

Theoretical SSCCs were compared with the experimental constants using nine Asp residues (Asp28, Asp34, Asp51, Asp69, Asp70, Asp122, Asp127, Asp136, and Asp143) obtaining the side chain torsion angle χ_1. The atypical values obtained for two residues (Asp62 and Asp63) that are far from showing a defined conformation, are due to the fact that considering three alternate rotamers none of them has a population value low enough to be neglected. This result could reflect a more complex dynamics for these two residues and show a possible increased mobility in this part of the protein. Theoretical results of the value of χ_1 at the established calculation level confirm experimental data obtained by NMR and X-ray. The Fourier coefficients for the Asp residues (Table 2) can be used to generate theoretical SCCCs for any Asp-type protein residue and predict the side chain angles.

AUTHOR CONTRIBUTIONS

María Botella: Investigation. **Ignacio Ema:** Investigation. **Jesús San Fabián:** Investigation. **Jose Manuel García de la Vega:** Investigation.

ACKNOWLEDGMENTS

Computer time provided by the Centro de Computación Científica of Universidad Autónoma de Madrid is gratefully acknowledged.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supplementary material of this article or from the corresponding author upon reasonable request.

ORCID

Ignacio Ema https://orcid.org/0000-0002-3003-213X

Jesús San Fabián https://orcid.org/0000-0002-3824-8277

Jose Manuel García de la Vega https://orcid.org/0000-0002-1940-422X

REFERENCES

[1] N. V. Bhagavan, C.-E. Ha, Three-dimensional structure of proteins. in *Essentials of Medical Biochemistry*, Elsevier Academic Press, London **2011**, p. 29. https://doi.org/10.1016/b978-0-12-095461-2.00004-7

[2] A. S. Holehouse, R. Pappu, R. V. Collapse, *Annu. Rev. Biophys.* **2018**, *47*, 19.

[3] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, W. H. Freeman, New York **2002**.

[4] R. B. Best, J. Clarke, M. K. Karplus, *J. Am. Chem. Soc.* **2004**, *126*, 7734.

[5] V. Muñoz, *Protein Folding, Misfolding and Aggregation: Classical Themes and Novel Approaches*, RSC Publishing, Cambridge, UK **2008**.

[6] M. P. Williamson, T. F. Havel, K. Wüthrich, *J. Mol. Biol.* **1985**, *182*, 295.

[7] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, M. Rance, N. J. Skelton, *Protein NMR Spectroscopy. Principles and Practice*, Elsevier Academic Press, Burlington **2007**.

[8] J. R. Allison, W. F. van Gunsteren, *Chem. Phys. Chem.* **2009**, *10*, 3213.

[9] D. Steiner, J. R. Allison, A. P. Eichenberger, W. F. van Gunsteren, *J. Biomol. NMR* **2012**, *53*, 223.

[10] H. Tandel, K. Florence, A. Misra, 9 - protein and peptide delivery through respiratory pathway. in *Challenges in Delivery of Therapeutic Genomics and Proteomics*, Elsevier, London **2011**, p. 429. https://doi.org/10.1016/B978-0-12-384964-9.00009-8

[11] J. Sancho, *Cell. Mol. Life Sci.* **2006**, *63*, 855.

[12] L. A. Inda, M. L. Peleato, *Phytochemistry* **2003**, *63*, 303.

[13] Y. Astuti, E. Topoglidis, P. B. Briscoe, A. Fantuzzi, G. Gilardi, J. R. Durrant, *J. Am. Chem. Soc.* **2004**, *126*, 8001.

[14] L. B. Krivdin, *Molecules* **2021**, *26*, 2450.

[15] R. Suardíaz, R. Crespo-Otero, C. Pérez, J. San Fabián, J. M. G. de la Vega, *J. Chem. Phys.* **2011**, *134*, 061101.

[16] P. Salvador, *Annu. Rep. NMR Spectrosc.* **2014**, *81*, 185.

[17] J. San Fabián, S. Omar, J. M. G. de la Vega, *Brief Bioinform.* **2021**, *22*, bbab020.

[18] J. San Fabián, I. Ema, S. Omar, J. M. G. de la Vega, *J. Chem. Inf. Model.* **2021**, *61*, 6012.

[19] M. Karplus, *J. Chem. Phys.* **1959**, *30*, 11.

[20] M. Karplus, *J. Am. Chem. Soc.* **1963**, *85*, 2870.

[21] C. A. G. Haasnoot, F. A. A. M. De Leeuw, C. Altona, *Tetrahedron* **1980**, *36*, 2783.

[22] K. G. R. Pachler, *Tetrahedron Lett.* **1970**, *22*, 1955.

[23] J. S. Fabián, S. Omar, J. M. G. de la Vega, *J. Chem. Theory Comput.* **2019**, *15*, 4252.

[24] J. Hermans, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3095.

[25] Y. Ching-Hsing, M. A. Norman, L. Schäfer, M. Ramek, A. Peeters, C. van Alsenoy, *J. Mol. Struct.* **2001**, *567-568*, 361.

[26] S. C. Lovell, I. W. Davis, W. B. Arendall III, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, D. C. Richardson, *Proteins: Struct., Funct., Genet.* **2003**, *50*, 437.

[27] E. Díez, J. San Fabián, J. Guilleme, C. Altona, L. A. Donders, *Mol. Phys.* **1989**, *68*, 49.

[28] C. Pérez, F. Löhr, H. Rüterjans, J. M. Schmidt, *J. Am. Chem. Soc.* **2001**, *123*, 7081.

[29] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.

[30] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.

[31] P. C. Hariharan, J. A. Pople, *Chem. Phys. Lett.* **1972**, *16*, 217.

[32] P. C. Hariharan, J. A. Pople, *Theor. Chem. Acc.* **1973**, *28*, 213.

[33] M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, J. A. Pople, *J. Chem. Phys.* **1982**, *77*, 3654.

[34] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09*, *Revision D.01*, Gaussian, Inc, Wallingford, CT **2013**.

[35] P. J. Wilson, T. J. Bradley, D. J. Tozer, *J. Chem. Phys.* **2001**, *115*, 9233.

[36] P. F. Provasi, G. A. Aucar, S. P. A. Sauer, *J. Chem. Phys.* **2001**, *115*, 1324.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: M. Botella, I. Ema, J. San Fabián, J. M. García de la Vega, *Int. J. Quantum Chem.* **2022**, e26979. https://doi.org/10.1002/qua.26979