STRUCTURAL BIOLOGY
ISSN 2059-7983

A radical approach to radicals

Youjia Liu,[a] Malgorzata Biczysko[a] and Nigel W. Moriarty[b]*

[a]International Center for Quantum and Molecular Structures, Shanghai University, Shanghai 200444, People's Republic of China, and [b]Molecular Biosciences and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720-8235, USA. *Correspondence e-mail: nwmoriarty@lbl.gov

Received 16 June 2021
Accepted 18 October 2021

Edited by R. J. Read, University of Cambridge, United Kingdom

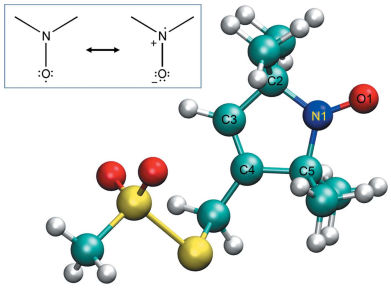
Keywords: ligand restraints; macromolecular refinement; quantum chemistry; spin labelling; nitroxides; chemical radicals.

Supporting information: this article has supporting information at journals.iucr.org/d

Nitroxide radicals are characterized by a long-lived spin-unpaired electronic ground state and are strongly sensitive to their chemical surroundings. Combined with electron paramagnetic resonance spectroscopy, these electronic features have led to the widespread application of nitroxide derivatives as spin labels for use in studying protein structure and dynamics. Site-directed spin labelling requires the incorporation of nitroxides into the protein structure, leading to a new protein–ligand molecular model. However, in protein crystallographic refinement nitroxides are highly unusual molecules with an atypical chemical composition. Because macromolecular crystallography is almost entirely agnostic to chemical radicals, their structural information is generally less accurate or even erroneous. In this work, proteins that contain an example of a radical compound (Chemical Component Dictionary ID MTN) from the nitroxide family were re-refined by defining its ideal structural parameters based on quantum-chemical calculations. The refinement results show that this procedure improves the MTN ligand geometries, while at the same time retaining higher agreement with experimental data.

1. Introduction

Of the more than 170 000 biological macromolecular structures in the Protein Data Bank (PDB; Burley *et al.*, 2019), over 31 000 (as of 6 April 2021) include at least one ligand. These ligands are listed in the Chemical Component Dictionary (CCD; Westbrook *et al.*, 2015). Many of them have a specific function in the organism that requires an additional chemical entity to the standard amino acids. These endogenous ligands can be studied using macromolecular crystallography to understand natural processes. Other natural or artificial ligands are more mobile, leading to the field of structure-driven drug design (SDDD): binding a molecule (often synthesized) to a protein for therapeutic purposes. X-ray crystallographic studies of these protein–ligand complexes are performed to determine the efficacy of drug candidates. These types of studies can be hampered by a third class of ligand: molecules that are required to conduct the X-ray crystallography experiment. These include cryoprotectants, crystallization conditions and lipids. These agents aim to help the experiment, but they can also hinder experimental outcomes by, for example, causing significant structural change or binding preferentially. Another ligand class includes molecules that are introduced to aid in other experimental techniques.



An interesting example of the last class is represented by the nitroxides family, which are aromatic radicals characterized by a long-lived spin-unpaired electronic ground state: the unpaired electron is localized on the N—O group (Jeschke, 2013; Torricella *et al.*, 2021). For this reason nitroxide radicals are often used in site-directed spin labelling (SDSL; Klare &

Steinhoff, 2009) of diamagnetic biomacromolecules (Kugele *et al.*, 2019), which in turn allows the study of the stereochemical structure and dynamical information of biomolecular complexes using electron paramagnetic resonance[1] (EPR; Altenbach *et al.*, 1990; Zerbetto *et al.*, 2007; Torricella *et al.*, 2021). The latter is a spectroscopic technique based on magnetic resonance principles that allows the characterization of systems containing unpaired electrons to determine information on their chemical environment with high resolution. Modern spectroscopic techniques such as high-field EPR and electron–nuclear double resonance (ENDOR) provide abundant information about the interaction of unpaired electrons with the surrounding magnetic nuclei encoded in hyperfine (A) and gyromagnetic (g) tensors (Brustolon & Doorslaer, 2011). In the case of proteins, the EPR technology coupled with SDSL exploits the fact that the magnetic properties of nitroxides are strongly sensitive to their local environment, allowing the investigation of dynamics through line-shape analysis (Hubbell *et al.*, 1996; Klare, 2013) or distance determination (Jeschke, 2012).

Nitroxide radicals can be classified into two groups: those with a five-membered and those with a six-membered aromatic ring, with the former usually having a larger nitrogen spin density due to the planar geometry favouring an electronic resonance structure (Improta & Barone, 2004). Several computational and experimental studies indicate that the N—O moiety is planar in the five-membered ring structures and pyramidal in the six-membered ring structures (Improta & Barone, 2004). One of the most commonly used spin labels is *S*-[(1-oxyl-2,2,5,5-tetramethyl-pyrrol-3-yl)methyl] methanesulfonothioate (CCD ID MTN), which contains a five-membered ring (Fig. 1). In analogy to the smaller five-membered representative of the nitroxides family, PROXYL (2,2,5,5-tetramethyl-1-pyrrolidinyloxy), it is expected to have a planar conformation of the nitroxide moiety (Pedone *et al.*, 2010). The staggered conformation between the O atom and the substituents on the adjacent C atoms minimizes the steric clashes in the planar composition of these five-membered cycles (Improta & Barone, 2004). The resonance structures of the radical moiety are shown in the inset in Fig. 1. Note the presence of the unpaired electron denoted by the 'dot' near the N atom. Other examples of nitroxide radicals, such as TEMPO (2,2,6,6-tetramethyl-1-piperidin-*N*-oxyl), PROXYL, TOAC (2,2,6,6-tetramethyl-*N*-oxyl-4-amino-4-carboxylic acid) and others (Stendardo *et al.*, 2010), could be used as possible precursors for ferromagnetic materials (Chiarelli *et al.*, 1993), which are employed in the biomedical field to study cell membrane fluidity or utilized to enhance flexible energy-storage device abilities (Xie *et al.*, 2019).

EPR studies of spin-labelled proteins can be connected with crystallographic experiments, allowing the determination of a static atomistic description of the whole molecular model through a structure-refinement procedure (Akter *et al.*, 2019). By comparing EPR spectra taken prior to and after irradiation, it has been demonstrated that the radical nature of nitroxide can be preserved during an X-ray experiment (Consentius *et al.*, 2016). The refinement requires additional *a priori* information to compensate for the lack of high-resolution data and the typically poor ratio of experimental observations to model parameters in refinement. These are supplied to the refinement in the form of chemical restraints.

Chemical restraints should be at least as accurate as the experimental data demands. Accurate ideal values and estimated standard deviation (e.s.d.) values for restraints can be obtained from the small-molecule experiments archived in the Cambridge Structural Database (CSD; Groom *et al.*, 2016; Taylor & Wood, 2019) and the Crystallography Open Database (COD; Vaitkus *et al.*, 2021) or from high-resolution ligand-containing macromolecular structures in the PDB. Unfortunately, the restraints used in the refinement play a role in the final result deposited in the PDB. This is particularly troublesome for entities that occur in low numbers or when all of the entries used poor restraints. This requires validation-based filtering to remove biased entries. The CSD is less prone to this bias and the geometry information is easily accessible to the user or programmer via the *Mogul* interface (Bruno *et al.*, 2004; Cottrell *et al.*, 2012).

Alternatively, ligand restraints can be based on quantum-mechanical (QM) computations. Drug-like molecule geometries obtained from semi-empirical QM methods are within the accuracy required for the majority of situations. An even simpler (for the user) alternative involves parameterizing the 'types' of bonds and angles found in a molecule, so that any new molecular structure can be constructed based on lookup tables.

Unfortunately, there are several situations in which unusual chemical bonding prevents the use of such generic databases or simplified QM models, requiring alternative schemes for obtaining restraints. Metal clusters require a detailed study of the small-molecule databases to obtain accurate restraints (Moriarty & Adams, 2019). Molecules containing a single metal ion are generally outside the range of semi-empirical

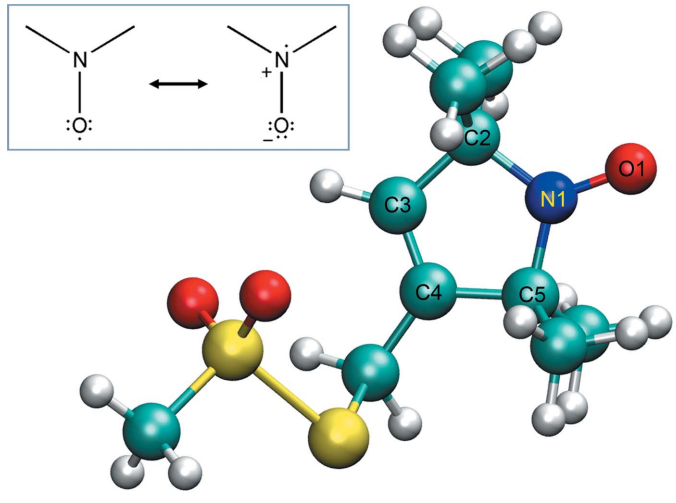


Figure 1
Representations of the complete MTN radical as referenced in the CCD. Inset: schematic representation of radical moiety resonance structures.

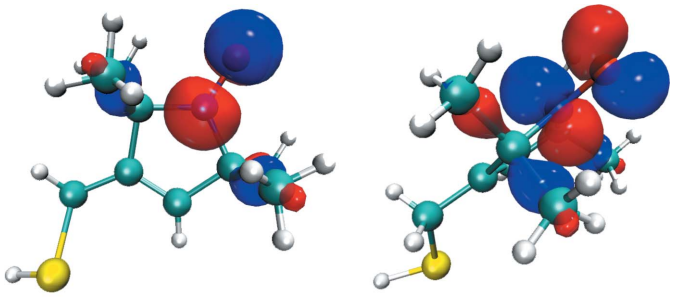
[1] Alternatively known as electron spin resonance (ESR).

QM and can also be problematic for the restraint-generation software due to the more complex nature of metal coordination. Purely covalently bonded ligands with typical bonding patterns (for example, drug-like molecules) are candidates for semi-empirical QM, *Mogul* or lookup tables.

One emerging class of molecule with covalent bonds that is not completely integrated into the biomacromolecular structure field (resources and software) is the chemical radical. The central resources for ligand information in the PDB websites are not static; they are always improving. For instance, at the start of this work the Chemical Components Database (CCD; Westbrook *et al.*, 2015) for MTN had an H atom bound to the O atom. This has changed, along with the visual representation of the ligand, to include the radical 'dot' on the RCSB website (Berman *et al.*, 2000), albeit on the wrong atom. However, in reality the nitroxide radical moiety is best described by a resonance structure (Fig. 1, inset), with the single occupied molecular orbital (SOMO; Fig. 2) corresponding to the antibonding π^* orbital localized on both the N and O atoms of the nitroxide. The contribution of the N atom's atomic orbital to the SOMO provides an estimate of the unpaired electron spin density on the N atom, while the final contribution to the radical character from either of the atoms is determined by the subtle interplay of stereo-electronic effects (Improta & Barone, 2004).

Furthermore, the SMILES string and the geometry of the N atom are still lacking the radical information. The standard SMILES definition (Weininger, 1988) does not explicitly specify radicals. It requires the explicit specification of H atoms or a nonstandard extension that will require firstly adoption by the PDB and secondly the likely retrofitting of restraint-generation software. The geometry of the radical moiety is greatly affected by the inclusion of the erroneous H atom or any misunderstanding of the planar nature of the atoms around the N atom. It should be noted that MTN is neutral.

A further nuance of this entity is related to the leaving group upon covalently binding to the S^{γ} atom of a cysteine residue. This leaving group is relatively large (SO_2CH_3), thus complicating the diagrams in PDBe (Velankar *et al.*, 2016), but also reduces the number of restraint values in MTN that need to be accurate. This greatly simplifies interrogation of the CSD and increases the statistical significance.



(*a*) (*b*)

Figure 2
Single-occupied molecular orbital (SOMO) computed for the MTN structure extracted from re-refined PDB entry 6n87; perpendicular views (*a*) and (*b*).

Table 1
PDB codes and resolutions of all proteins considered in this study.

The proteins are divided into two groups: high resolution (better than 2 Å) and low resolution (2 Å and worse).

PDB entry	Resolution (Å)
5jdt	1.0
2igc	1.4
2xiu	1.5
6n87	1.59
5bmh	1.6
5g27	1.61
2nth	1.8
4ek1	1.97
6pgy	2.0
6pgz	2.0
2q9e	2.1
4wwl	2.23
3mpn	2.25
3mpq	2.25
2xga	2.3
3stz	2.5
5bmi	2.5

2. Methods

2.1. Selection of protein models

From a total of 40 protein structures in the Protein Data Bank (PDB; Burley *et al.*, 2019) containing MTN (as of 6 April 2021), six NMR structure entries were excluded. For the remaining 34 entries, we calculated the stereochemical information for a protein structure from the atomic coordinates deposited in the PDB and carefully screened the model quality and completeness based on the *MolProbity* scores (Chen *et al.*, 2010; Williams *et al.*, 2018) and detailed analysis of electron-density map fits. In particular, OMIT polder maps (Liebschner *et al.*, 2017) were computed in order to select proteins with well defined and complete electron density on the MTN ligands. The final selected 17 PDB entries (containing 26 MTN instances) are listed in Table 1, along with the corresponding experimental resolutions. Excluded models and criteria are listed in Supplementary Table S1.

2.2. Preparation of the MTN ligand restraints file

In order to more accurately describe the properties of aromatic nitroxides in terms of well defined physical-chemical effects, flexibility and delocalization both need to be considered. For QM computations, we employed the density functional theory (DFT) with parameter-free PBE0 (Adamo & Barone, 1999) functional, which includes an amount of exact exchange, and the polarized double-ζ basis set N07D, which is purposely tailored for DFT-based calculations of the structural, dynamic and magnetic properties of organic free radicals, including those from the nitroxide family (Barone *et al.*, 2008; Barone & Cimino, 2009; Pedone *et al.*, 2010; Pavone *et al.*, 2010). Moreover, dispersion effects have been included by means of Grimme's dispersion correction D3 (Grimme *et al.*, 2010) in conjunction with Becke–Johnson damping (Grimme *et al.*, 2011). The structural features of MTN (bond lengths, bond angles and dihedral angles) were obtained by unconstrained geometry optimization, with very tight convergence

Table 2
Selected geometry values for the ring structure in MTN from the Cambridge Structural Database and QM geometry minimization described in the text. The values are averaged based on the symmetry of the ring.

Bond lengths are in ångströms and angles are in degrees.

	Mean	Standard deviation	QM
N—O	1.272	0.006	1.258
N—C	1.480	0.006	1.479/1.477
C—C	1.505	0.011	1.515/1.499
C═C	1.326	0.007	1.338
O—N—C	122.3	0.5	122.4/122.7
C—N—C	115.2	0.4	114.9
N—C—C	99.3	0.5	100.0
C—C═C	112.9	0.8	111.8/113.2
O—N—C—C	180.0	1.7	178.8

criteria, which was followed by computation of harmonic frequencies. The latter allowed confirmation of the nature of the stationary point as a minimum structure and identification of the nuclear motions related to the low-frequency vibrations. All calculations were performed with *Gaussian*16 (Frisch *et al.*, 2016) on the complete MTN entity as described in the CCD (Fig. 1). The ideal values for the internal coordinates of the ring structure of MTN are given in Table 2. The last column lists the values from the QM geometry minimization. Interestingly, approximately symmetrically equivalent bonds and angles can be uniquely denoted by using the atomic element and bond order (using a single bond for the radical). There is only one N—O bond or C═C bond. There are two almost equivalent C—C bonds, which are both given in the QM results.

To confirm the accuracy of the QM results, a search of the CSD was performed. The query was constructed as shown in Supplementary Fig. S1. This resulted in 94 examples of the moiety that were filtered for outliers using Tukey's fences (Beyer, 1981). The remaining 71 instances produced the geometric values shown in Table 2. Symmetrically equivalent internal coordinates are averaged for the CSD values. The QM and CSD values are in close agreement. Of interest are the CSD values for the O—N—C—C torsion, which are completely planar. Furthermore, the standard deviation (1.7°) covers the QM value of 178.8°, which is quite flexible.

The flexibility of the radical moiety prompted an investigation of the low-frequency vibrations. The frequencies were calculated using the same QM method and basis set and were visualized for their effect on the moiety. The focus was on the low frequency vibration involving out-of-plane deformations within the nitroxide moiety, which is predicted at about 120 cm^{-1}. The potential energy scan along the out-of-plane N—O bending coordinate was computed to define the range of displacements accessible at the zero-point vibrational level (ZPVE; Fig. 3). The scan showed that the O atom can bend out of plane by up to ±0.12 Å, which corresponds to an O—N—C—C torsion range of ±8°. The usual e.s.d. value for planar restraints is 0.02 Å, but a feature of the planarity restraint in *Phenix* that allows individual e.s.d. values for each atom has proven to be effective in the case of arginine (Moriarty *et al.*, 2020). This feature was used to allow more flexibility of the O atom while maintaining the general planarity of the ring. These QM-based restraints are available in GeoStd (N. W. Moriarty & P. D. Adams; http://sourceforge.net/projects/geostd) and are distributed with *Phenix* (Liebschner *et al.*, 2019).

The QM unconstrained geometry minimizations were used for the geometry data of MTN, including bond lengths, bond angles and dihedral angles, as well as the allowed range of N—O deviations from the molecular planes. The restraints file was generated by *eLBOW* (Moriarty *et al.*, 2009) and edited in *REEL* (Moriarty *et al.*, 2017).

2.3. Refinement protocols

Models containing the MTN ligand with experimental data in the PDB (Burley *et al.*, 2019) were obtained via the RCSB web server (Berman *et al.*, 2000). H atoms, which in reality constitute about 50% of the atoms in a macromolecular structure, were added using *ReadySet!* (*phenix.ready_set*). Beyond the addition of H atoms, *ReadySet!* generated other ligand restraints and metal-coordination restraints.

Each model was refined using *phenix.refine* (Afonine *et al.*, 2012) for ten macrocycles with weight optimization. The resulting models were evaluated with the suite of validation methods in *Phenix* including the metrics in *MolProbity* (Chen *et al.*, 2010; Williams *et al.*, 2018) and the recently implemented Rama-Z score (Sobolev *et al.*, 2020). Graphical programs such as *PyMOL* (version 1.8; Schrödinger) and *Coot* (Emsley *et al.*, 2010) can be used to visualize models and data files.

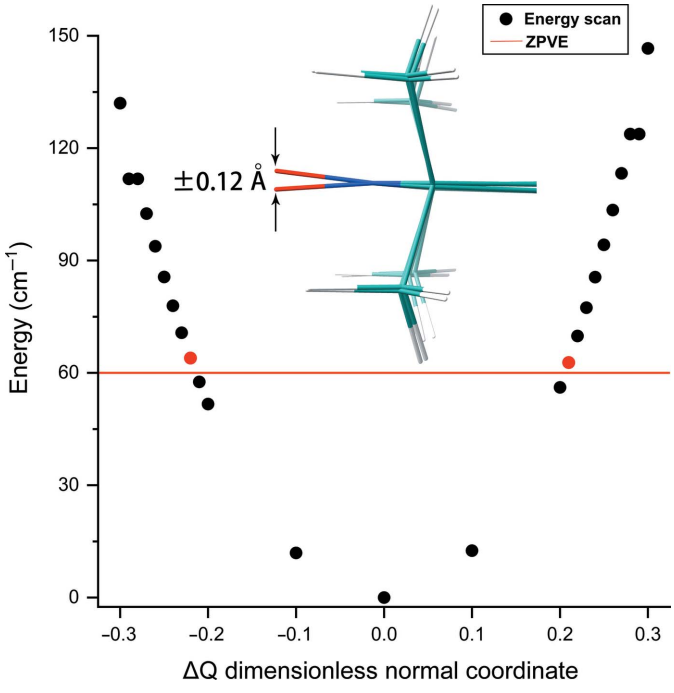


Figure 3
Scanning along the out-of-plane N—O bending normal coordinate of MTN allows definition of the flexibility of the nitroxide group. The two structures displaced along a normal mode with energy crossing the ZPVE (displacements of −0.22 and 0.21) are superposed and marked by red dots on a plot.

2.4. Electron localization function (ELF)

Introduced by Becke and Edgecombe, the electron localization function (ELF; Becke & Edgecombe, 1990) is a scalar field that ranges from 0 to 1 which reflects the probability of finding same-spin electrons close together. Topological analysis of the ELF allows partition of the molecular space into core and valence attractor basins, for which properties such as an average electron population can be calculated. Topological analyses were performed using *Multiwfn* (Lu & Chen, 2012) for a parallel epipedic grid of points with a step of 0.10 bohr based on the results of PBE0/N07D computations. The graphical representation of ELF isosurfaces was constructed with the help of *UCSF Chimera* (Pettersen *et al.*, 2008).

ELF can be calculated separately for alpha and beta electrons, which for closed-shell systems gives equivalent patterns of electron localization. For the open-shell systems here we analyze the ELF calculated for the alpha-electron population, *i.e.* that with one more electron. This approach has proved successful in analyzing the bonding situation in radical systems (Bil & Latajka, 2004*a*,*b*, 2005; Bil *et al.*, 2007, 2013).

PBE0/N07D computations have also been used to obtain atomic charges based on full natural bond orbital (NBO; Foster & Weinhold, 1980; Reed *et al.*, 1988) analysis as implemented in *Gaussian*16 (Frisch *et al.*, 2016).

2.5. Noncovalent interactions

Chemical interactions between a protein and a drug molecule are dominated by noncovalent interactions (Johnson *et al.*, 2010). This is also true for covalently bound ligands such as MTN, for which the weak interactions define the final position in the protein pocket. For the analysis of weak intermolecular interactions between a protein and an MTN ligand, the noncovalent interaction (NCI) method based on the electron density and its derivatives has been employed.

Noncovalent interactions are characterized by a reduced density gradient (RDG) function, which can be located by generating gradient surfaces enclosing the corresponding regions of real space. The interaction types can be discriminated by RDG isosurfaces based on the sign of the second density Hessian eigenvalue λ_2 and the density (ρ) itself in these regions. These isosurfaces were visualized by *VMD* (Humphrey *et al.*, 1996).

The NCI computations were performed for the protein fragment surrounding the MTN ligand. A suitable part of the protein was selected based on the clustering procedure developed within the *Q*|*R* (quantum-refinement) project (Zheng, Reimers *et al.*, 2017) that enables the refinement of entire protein models using QM gradients. One technique necessary for *Q*|*R* is the ability to break models into smaller clusters that are chemically meaningful (Zheng, Moriarty *et al.*, 2017). For the analysis of noncovalent interactions, we have to consider all aspects of the whole molecular model including all H atoms on solvent (water) molecules, which is performed by *ReadySet!* (`add_h_to_water=True`). The clustering procedure developed in *Q*|*R*, which itself uses NCI analysis, was applied to create smaller molecular fragments that are amenable to NCI computations. This procedure generated a molecular model containing MTN and its surroundings (391 atoms: 20 residues and seven waters). The NCI analysis usually employs density from QM computations; however, calculation of the real DFT density for such a large system would be extremely expensive. An alternative possibility employed in protein studies is to use the promolecular density (Spackman & Maslen, 1986; Johnson *et al.*, 2010) obtained by simply summing exponential atomic densities. We have tested that such a procedure can be employed for radicals by comparing the NCI analysis from DFT computations and promolecular densities for an isolated MTN radical; both shown qualitatively identical results. Hence, we use the *NCIplot* (Contreras-García *et al.*, 2011) protein–ligand interaction code employing the approximated promolecular density to calculate the gradient isosurfaces.

3. Results

3.1. Comparison of initial and refined models

We start by comparing the overall global crystallographic and model statistic metrics between the initial structure (deposited in the PDB) and the re-refined structure. These comparisons are not specifically directed at the changes in the MTN ligand caused by the improved restraints, but are rather a sanity check of the initial models and the final models from the refinements. Similar or improved *R* factors are expected upon re-refinement.

As expected, there is an overall improvement in both *R* factors ($R_{\rm work}$ and $R_{\rm free}$) as well as in $R_{\rm gap}$. None of the initial *R* factors are suspicious. Furthermore, the refined structures have smaller bond and angle root-mean-square deviation (r.m.s.d) values of below 0.016 Å and 1.7°, respectively, compared with largest deviations of 0.127 Å and 2.4° for the initial models. The other validation parameters are not altered by a significant amount, as shown in Supplementary Table S2. Moreover, 14 of the refined structures show a Rama-Z score below 2 (reliable backbone distributions), while all initial and refined structures are within the range of possible geometries, with an absolute *Z*-score of lower than 3. In conclusion, the re-refinements have led to some small improvements in the overall molecular models, so the new structures are at least as good as the initial structures deposited in the PDB.

3.2. Comparison of MTN ligand structure parameters

Focusing on the stereochemical structure of MTN, the bond lengths, bond angles and dihedral angles are reported in Fig. 4. All values are shown in Supplementary Table S3. The protein models are divided into two groups by resolution, with the boundary being between 1.97 and 2.0 Å resolution (Table 1). The plots in the left-hand column of the figure are for the lower resolution structures, with the higher resolution structures on the right. Each plot has up to two geometric features, comparing the initial absolute deviation from ideal values presented in this work using a pink box plot with whiskers (BPW) with the final BPW values in cyan. The most illus-

trative bond lengths, N—O and N—C (averaged per MTN instance), are shown in Figs. 4(*a*) and 4(*b*). It is expected that the range of deviations will be smaller for the re-refined structures for both higher and lower resolution sets. The reduction between the initial and the re-refined structure is particularly large for the N—O bond, with initial deviations as large as 0.205 Å and final deviations smaller than 0.026 Å. This is likely to be due to misunderstanding of the nature of the radical moiety, leading to the use of a wide range of ideal bond lengths in the various deposited refinements.

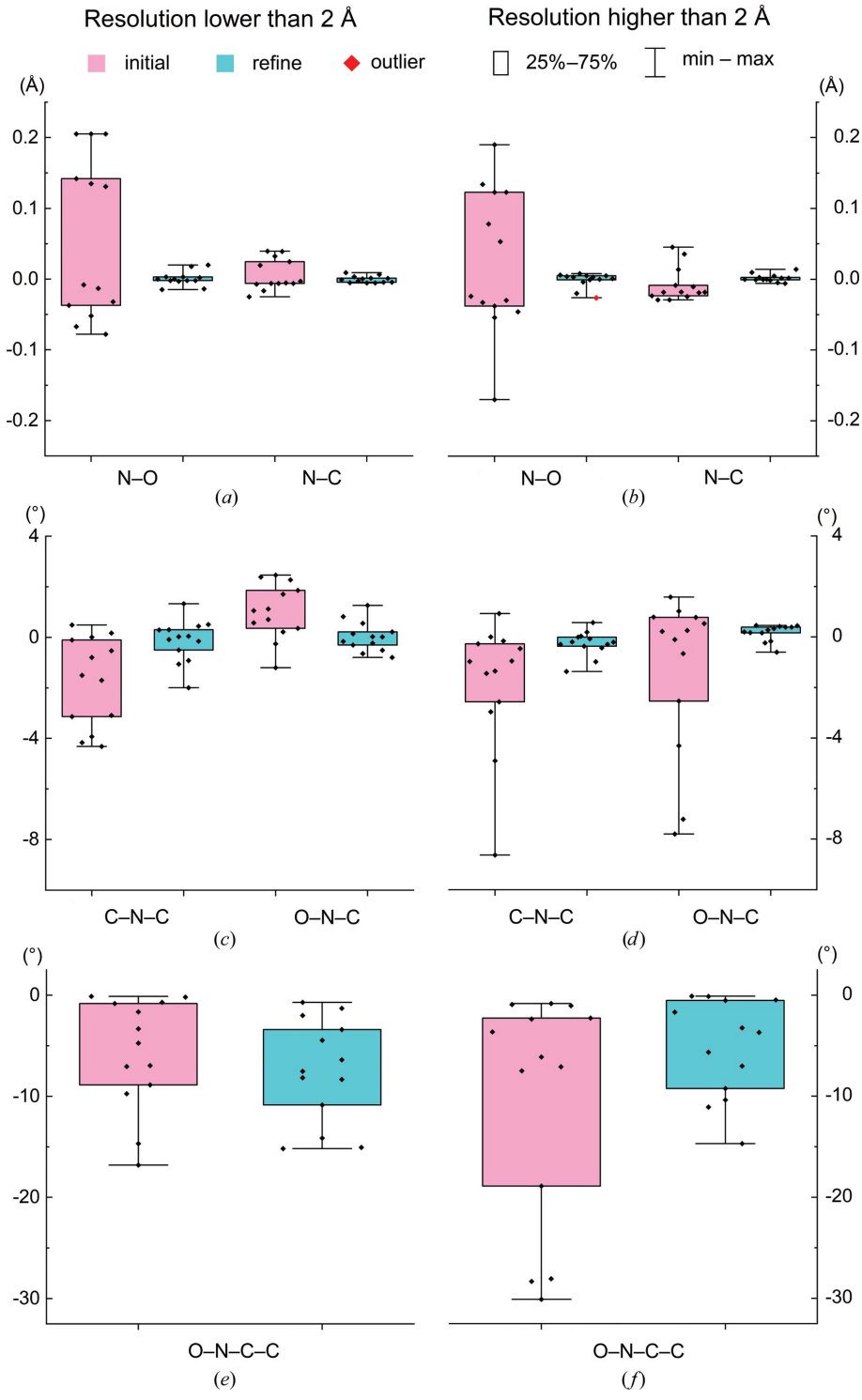


Figure 4
Deviations in bond length (*a*, *b*), bond angle (*c*, *d*) and dihedral angle (*e*, *f*). The resolution in the left panel is lower than 2 Å and that in the right panel is higher than 2 Å. The differences with respect to a reference (ideal value) are computed as Δinitial/final (pink/cyan), which corresponds to the initial/refined value minus the ideal value.

The representative angles O—N—C and C—N—C have their deviations from ideal presented in Figs. 4(*c*) and 4(*d*). The range of deviations is larger for the initial models, highlighting the range of values used in the restraints. Counterintuitively, for the re-refined models the range of deviation values is larger for the models at resolutions worse than 2 Å than for those at better resolutions. The higher resolution structures are more likely to diverge from the restraints based on the increased experimental information. However, in the limited sample in this study it appears that the experimental data are so poor that the ring structure is being compressed to fit the density blob. It should be also noted that the possibility of radiation damage, which would add some noise to the analysis, can not be fully excluded. The final range of deviations is also a reflection of using a relatively relaxed 3° for the e.s.d. value for all angles. No angle deviations exceed the e.s.d. value.

The increased flexibility of the plane is reflected by the increased variation of the torsion-angle deviations of the O—N—C—C dihedral (Figs. 4*e* and 4*f*). The deviations are smaller for the refined structures. Any pyramidal N atoms in the initial structures would reduce this difference, but the O atom can still be out of plane.

3.3. Comparison of the initial and refined MTN ligands with their electron-density maps

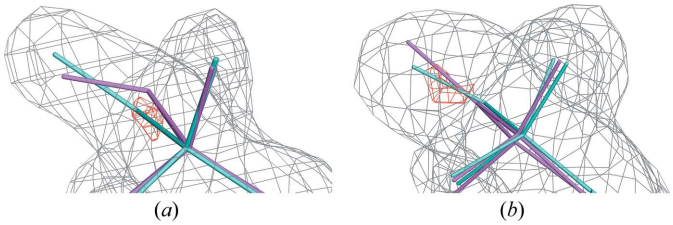
A more detailed analysis of the MTN molecular structures showing the largest differences upon refinement is shown in Fig. 5. The initially most 'bent' MTN from *Plasmodium falciparum* FVO apical membrane antigen (PDB entry 6n87) is shown in Fig. 5(*a*) along with the experimental electron density. The initial structure shows significant pyramidalization on the N atom, along with an out-of-plane deviation of the O atom and a chemically unjustified zigzag pattern of the O—N—C—C dihedral of

about 150°. After refinement, the N—O group becomes essentially coplanar (O—N—C—C = 179.9°) with respect to the aromatic ring. The experimental density indicates that the planar configuration is more likely.

In contrast, the MTN structure incorporated into the labelled T4 lysozyme pseudo-wild-type V75C mutant is reported as planar (O—N—C—C = 178.3°) in the deposited molecular model (PDB entry 6pgz), but the re-refined structure has an out-of-plane deformation of the O atom with an O—N—C—C dihedral of 164.8° (Fig. 5*b*). Interestingly, the adjacent dihedral angle in the ring (N—C—C═C) remains essentially planar. Once again, the initial geometry was restricted by overly strict planar restraints, as the experimental density clearly agrees with the bent geometry.

3.4. Electron localization function analysis

The radical character of MTN is illustrated by the localization domains of the electron localization function calculated for the initial and re-refined ligand structures extracted from PDB entries 6n87 and 6pgz. The basins for PDB entry 6n87 are shown in Fig. 6. The numerical results for the electron populations of ELF basins relevant to the N—O moiety are reported in Table 3. The remaining data for these structures are reported in Supplementary Table S4. As expected, most of the unpaired electron density is localized on the N—O unit within four basins representing the nonbonding electron density of the N and O atoms. Both initial structures are characterized by some degree of pyramidalization on the N atom, which is larger for PDB entry 6n87 than for PDB entry 6pgz, with N—C—C═C torsion angles of about 166.6° and 175.4°, respectively. Both refined geometries are almost planar, as reflected by N—C—C═C torsion angles of 179.9° and 177.7°, respectively. These modifications are reflected in the changes of the electron population on nitrogen, which is larger for PDB entry 6n87. Interestingly, for both refined MTN structures the total electron population in nitrogen basins [$V_{i=1,2}$(N)] is about 1.9–2.0 electrons. The refinement increases the electron population on oxygen for PDB entry 6n87 and decreases it for PDB entry 6pgz, so for the refined structures the two basins on the O atom, $V_{i=1,2}$(O) (Fig. 6), are populated with 2.9 electrons each; three electron pairs in total. In the basin representing the covalent N—O bond, the electron population decreases for PDB entry 6n87 and increases for PDB entry 6pgz, but the final result is the same for both refined geometries. The 1.2 electrons in V(N,O) suggest a single character of the N—O bond. Overall, the ELF analysis indicates that the increased planarity on the N atom is associated with an electron-density distribution consistent with the resonance structure presented on the right in the inset in Fig. 1 with the unpaired electron localized on the N atom.



(*a*) (*b*)

Figure 5
Electron-density maps of MTN ligands from PDB entries 6n87 and 6pgz; the initial structure is in pink and the refined structure is in cyan. (*a*) PDB entry 6n87: grey, $2mF_{obs} - DF_{model}$ map contoured at 1δ; red, $2mF_{obs} - DF_{model}$ map contoured at 2.4δ. (*b*) PDB entry 6pgz: grey, $2mF_{obs} - DF_{model}$ map contoured at 1δ; red, $2mF_{obs} - DF_{model}$ map contoured at 4.50δ.

Table 3
Electron populations of ELF basins and NBO charges relevant to the nitroxide moiety of MTN in PDB entries 6n87 and 6pgz.

	Electronic population	
Basin	6n87 (initial/refined)	6pgz (initial/refined)
V_1(O)	2.69/2.91	3.03/2.87
V_2(O)	2.69/2.90	3.04/2.91
V_1(N)	1.65/1.31	1.12/1.27
V_2(N)	†/0.60	1.08/0.67
V(N,O)	1.87/1.23	0.78/1.25

	Charge	
Atom	6n87 (initial/refined)	6pgz (initial/refined)
O	−0.39/−0.44	−0.42/−0.45
N	0.12/0.05	−0.03/0.04

† PDN entry 6n87 does not show an ELF basin of the V_2(N) type.

3.5. Noncovalent interactions

From these two example models, PDB entry 6n87 shows the interactions between the MTN ligands and the protein. The noncovalent interactions between molecules dominate the mutual recognition and specific binding between biological macromolecules, including various physiological processes and functions, and are also important in the field of drug design. The MTN is covalently bound to the protein by the disulfide bridge, but its final orientation inside the protein pocket is strongly dependent on weak interactions. Fig. 7

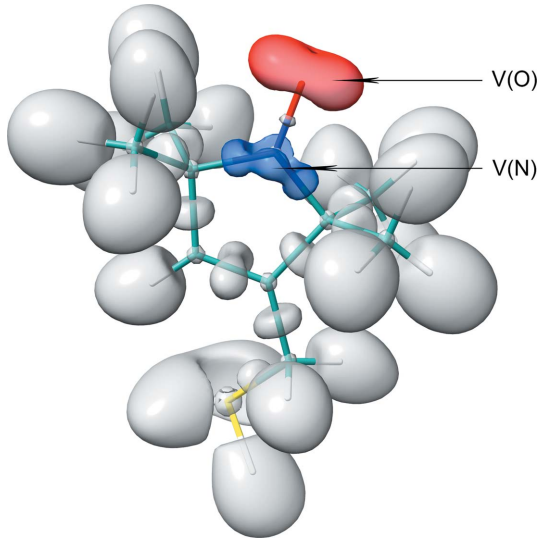


Figure 6
Localization domains of the ELF plotted for the refined MTN structure from PDB entry 6n87 using an η = 0.82 isosurface.

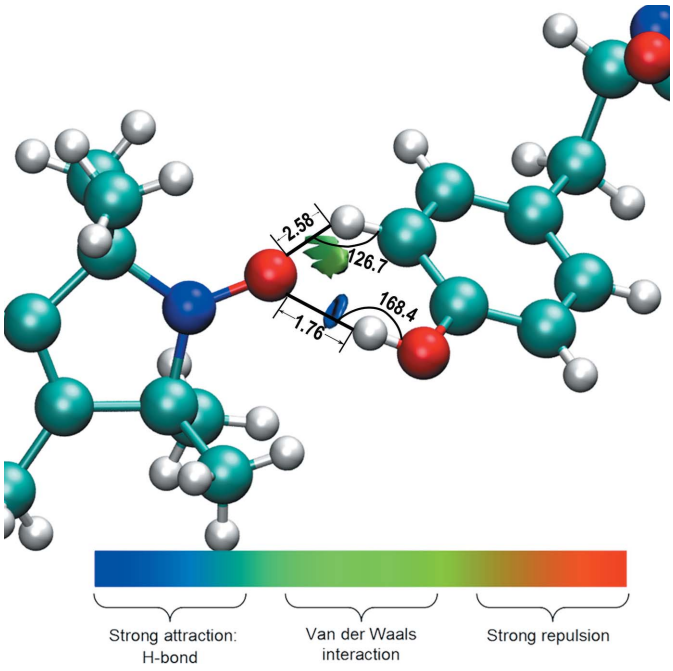


Figure 7
Noncovalent interactions of an MTN ligand with its surroundings (Tyr251) together with a schematic description of the interactions. The NCI RDG isosurface is shown at 0.3 arbitrary units.

shows the noncovalent interactions between the MTN ligand and its surroundings. These mainly correspond to very weak van der Waals-type attractive interactions along with some repulsion. However, we also note the hydrogen bonding involving the nitroxide part. The terminal O atom forms a hydrogen bond to the H^{η} atom of Tyr251. This O···H—O hydrogen bond is characterized by a 1.76 Å O···H distance, a 2.58 Å O···O distance and a 168.4° O···H—O angle, and thus can be considered to be a relatively strong hydrogen bond. This is consistent with the negative charge located on the O atom (Table 3). The possibility of a well defined hydrogen-bonding structure supports the modified geometry of the radical. NCI also reveals additional attraction between the N—O and C—H groups of Tyr251, which is characterized by H···O and C···O distances of 2.58 and 3.32 Å, respectively, and an O···H—C angle 126.7°, in line with a less directional van der Waals character.

For the re-refined MTN from PDB entry 6pgz, NCI analysis does not show any specific noncovalent interactions around the N—O part (an NCI plot isosurface of PDB entry 6pgz is shown in Supplementary Fig. S2).

4. Conclusions

We have shown that reliable restraints for the MTN ligand can be obtained based on quantum-chemical computations at the DFT level. MTN is a representative case of nitroxide radicals, which are commonly used as spin labels in EPR studies of proteins. The radical character of MTN is strongly associated with the N—O structural motif, with the excess electron density localized mostly as nonbonding electron density on the N atom. The radical character of MTN is often not recognized in reported PDB structures or crystallographic refinement because the H atoms are not seen by X-rays, which leads to some depositors not including them. This ambiguity then contributes to the terminal O atom of the N—O group being reported as capped by an H atom.

The DFT-derived restraints have been applied to re-refine 17 proteins including cases of the MTN radical with small and large distortions from planarity. In all cases re-refinement leads to a new molecular model with at least the same overall quality as the deposited structures. Detailed analysis of the refinement results show that the MTN ligand parameters are improved whilst retaining higher agreement with experimental electron densities. Additionally, an analysis of noncovalent interactions between the ligand and the protein surroundings shows that well defined hydrogen bonding supports the reliability of the refined structures.

The same quantum-chemical approach as applied in this work can be easily extended to define restraints for other ligands, including challenging ligands with radical character, which can support the refinement of molecular models containing *de novo*-developed spin labels.

Acknowledgements

The authors thank Dr P. V. Afonine, Dr M. P. Waller, Dr H. Kruse and Dr A. Bil for fruitful discussions.

Funding information

This work was enabled by financial support from the National Natural Science Foundation of China (grant No. 31870738).

References

Adamo, C. & Barone, V. (1999). *J. Chem. Phys.* **110**, 6158–6170.
Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Akter, M., Drinkwater, N., Devine, S. M., Drew, S. C., Krishnarjuna, B., Debono, C. O., Wang, G., Scanlon, M. J., Scammells, P. J., McGowan, S., MacRaild, C. A. & Norton, R. S. (2019). *ChemMedChem*, **14**, 603–612.
Altenbach, C., Marti, T., Khorana, H. G. & Hubbell, W. L. (1990). *Science*, **248**, 1088–1092.
Barone, V. & Cimino, P. (2009). *J. Chem. Theory Comput.* **5**, 192–199.
Barone, V., Cimino, P. & Stendardo, E. (2008). *J. Chem. Theory Comput.* **4**, 751–764.
Becke, A. D. & Edgecombe, K. E. (1990). *J. Chem. Phys.* **92**, 5397–5403.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Beyer, H. (1981). *Biom. J.* **23**, 413–414.
Bil, A., Berski, S. & Latajka, Z. (2007). *J. Chem. Inf. Model.* **47**, 1021–1030.
Bil, A., Grzechnik, K., Mierzwicki, K. & Mielke, Z. (2013). *J. Phys. Chem. A*, **117**, 8263–8273.
Bil, A. & Latajka, Z. (2004*a*). *Chem. Phys.* **303**, 43–53.
Bil, A. & Latajka, Z. (2004*b*). *Chem. Phys.* **305**, 243–252.
Bil, A. & Latajka, Z. (2005). *Chem. Phys. Lett.* **406**, 366–370.
Bruno, I. J., Cole, J. C., Kessler, M., Luo, J., Motherwell, W. D. S., Purkis, L. H., Smith, B. R., Taylor, R., Cooper, R. I., Harris, S. E. & Orpen, A. G. (2004). *J. Chem. Inf. Comput. Sci.* **44**, 2133–2144.
Brustolon, M. & Van Doorslaer, S. (2011). *Computational Strategies for Spectroscopy: From Small Molecules to Nano Systems*, edited by V. Barone, pp. 1–9. Hoboken: John Wiley & Sons.

Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C. X., Chen, L., Costanzo, L. D., Christie, C., Duarte, J. M., Dutta, S., Feng, Z. K., Ghosh, S., Goodsell, D. S., Green, R. K., Guranovic, V., Guzenko, D., Hudson, B. P., Liang, Y. H., Lowe, R., Peisach, E., Periskova, I., Randle, C., Rose, A., Sekharan, M., Shao, C. H., Tao, Y. P., Valasatava, Y., Voigt, M., Westbrook, J., Young, J., Zardecki, C., Zhuravleva, M., Kurisu, G., Nakamura, H., Kengaku, Y., Cho, H., Sato, J., Kim, J. Y., Ikegawa, Y., Nakagawa, A., Yamashita, R., Kudou, T., Bekker, G. J., Suzuki, H., Iwata, T., Yokochi, M., Kobayashi, N., Fujiwara, T., Velankar, S., Kleywegt, G. J., Anyango, S., Armstrong, D. R., Berrisford, J. M., Conroy, M. J., Dana, J. M., Deshpande, M., Gane, P., Gáborová, R., Gupta, D., Gutmanas, A., Koča, J., Mak, L., Mir, S., Mukhopadhyay, A., Nadzirin, N., Nair, S., Patwardhan, A., Paysan-Lafosse, T., Pravda, L., Salih, O., Sehnal, D., Varadi, M., Vařeková, R., Markley, J. L., Hoch, J. C., Romero, P. R., Baskaran, K., Maziuk, D., Ulrich, E. L., Wedell, J. R., Yao, H. Y., Livny, M. & Ioannidis, Y. E. (2019). *Nucleic Acids Res.* **47**, D520–D528.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Chiarelli, R., Novak, M. A., Rassat, A. & Tholence, J. L. (1993). *Nature*, **363**, 147–149.

Consentius, P., Gohlke, U., Loll, B., Alings, C., Müller, R., Heinemann, U., Kaupp, M., Wahl, M. & Risse, T. (2016). *J. Am. Chem. Soc.* **138**, 12868–12875.

Contreras-García, J., Johnson, E. R., Keinan, S., Chaudret, R., Piquemal, J.-P., Beratan, D. N. & Yang, W. (2011). *J. Chem. Theory Comput.* **7**, 625–632.

Cottrell, S. J., Olsson, T. S. G., Taylor, R., Cole, J. C. & Liebeschuetz, J. W. (2012). *J. Chem. Inf. Model.* **52**, 956–962.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Foster, J. P. & Weinhold, F. (1980). *J. Am. Chem. Soc.* **102**, 7211–7218.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, F., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A. Jr, Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B. & Fox, D. J. (2016). *Gaussian 16*. Gaussian Inc., Wallingford, Connecticut, USA.

Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. (2010). *J. Chem. Phys.* **132**, 154104.

Grimme, S., Ehrlich, S. & Goerigk, L. (2011). *J. Comput. Chem.* **32**, 1456–1465.

Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Cryst.* B**72**, 171–179.

Hubbell, W. L., Mchaourab, H. S., Altenbach, C. & Lietzow, M. A. (1996). *Structure*, **4**, 779–783.

Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.

Improta, R. & Barone, V. (2004). *Chem. Rev.* **104**, 1231–1254.

Jeschke, G. (2012). *Annu. Rev. Phys. Chem.* **63**, 419–446.

Jeschke, G. (2013). *Prog. Nucl. Magn. Reson. Spectrosc.* **72**, 42–60.

Johnson, E. R., Keinan, S., Mori-Sánchez, P., Contreras-García, J., Cohen, A. J. & Yang, W. (2010). *J. Am. Chem. Soc.* **132**, 6498–6506.

Klare, J. P. (2013). *Biol. Chem.* **394**, 1281–1300.

Klare, J. P. & Steinhoff, H. J. (2009). *Photosynth. Res.* **102**, 377–390.

Kugele, A., Silkenath, B., Langer, J., Wittmann, V. & Drescher, M. (2019). *ChemBioChem*, **20**, 2479–2484.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.

Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C. & Adams, P. D. (2017). *Acta Cryst.* D**73**, 148–157.

Lu, T. & Chen, F. (2012). *J. Theor. Comput. Chem.* **11**, 163–183.

Moriarty, N. W. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 16–20.

Moriarty, N. W., Draizen, E. J. & Adams, P. D. (2017). *Acta Cryst.* D**73**, 123–130.

Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Cryst.* D**65**, 1074–1080.

Moriarty, N. W., Liebschner, D., Tronrud, D. E. & Adams, P. D. (2020). *Acta Cryst.* D**76**, 1159–1166.

Pavone, M., Biczysko, M., Rega, N. & Barone, V. (2010). *J. Phys. Chem. B*, **114**, 11509–11514.

Pedone, A., Biczysko, M. & Barone, V. (2010). *ChemPhysChem*, **11**, 1812–1832.

Pettersen, E., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2008). *J. Comput. Chem.* **25**, 1605–1612.

Reed, A. E., Curtiss, L. A. & Weinhold, F. (1988). *Chem. Rev.* **88**, 899–926.

Sobolev, O. V., Afonine, P. V., Moriarty, N. W., Hekkelman, M. L., Joosten, R. P., Perrakis, A. & Adams, P. D. (2020). *Structure*, **28**, 1249–1258.

Spackman, M. A. & Maslen, E. N. (1986). *J. Phys. Chem.* **90**, 2020–2027.

Stendardo, E., Pedone, A., Cimino, P., Menziani, M. C., Crescenzi, O. & Barone, V. (2010). *Phys. Chem. Chem. Phys.* **12**, 11697–11709.

Taylor, R. & Wood, P. A. (2019). *Chem. Rev.* **119**, 9427–9477.

Torricella, F., Pierro, A., Mileo, E., Belle, V. & Bonucci, A. (2021). *Biochim. Biophys. Acta*, **1869**, 140653.

Vaitkus, A., Merkys, A. & Gražulis, S. (2021). *J. Appl. Cryst.* **54**, 661–672.

Velankar, S., van Ginkel, G., Alhroub, Y., Battle, G. M., Berrisford, J. M., Conroy, M. J., Dana, J. M., Gore, S. P., Gutmanas, A., Haslam, P., Hendrickx, P. M., Lagerstedt, I., Mir, S., Fernandez Montecelo, M. A., Mukhopadhyay, A., Oldfield, T. J., Patwardhan, A., Sanz-García, E., Sen, S., Slowley, R. A., Wainwright, M. E., Deshpande, M. S., Iudin, A., Sahni, G., Salavert Torres, J., Hirshberg, M., Mak, L., Nadzirin, N., Armstrong, D. R., Clark, A. R., Smart, O. S., Korir, P. K. & Kleywegt, G. J. (2016). *Nucleic Acids Res.* **44**, D385–D395.

Weininger, D. (1988). *J. Chem. Inf. Model.* **28**, 31–36.

Westbrook, J. D., Shao, C., Feng, Z., Zhuravleva, M., Velankar, S. & Young, J. (2015). *Bioinformatics*, **31**, 1274–1278.

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. & Richardson, J. S. (2018). *Protein Sci.* **27**, 293–315.

Xie, Y., Zhang, K., Monteiro, M. J. & Jia, Z. (2019). *Appl. Mater. Interfaces*, **11**, 7096–7103.

Zerbetto, M., Carlotto, S., Polimeno, A., Corvaja, C., Franco, L., Toniolo, C., Formaggio, F., Barone, V. & Cimino, P. (2007). *J. Phys. Chem. B*, **111**, 2668–2674.

Zheng, M., Moriarty, N. W., Xu, Y., Reimers, J. R., Afonine, P. V. & Waller, M. P. (2017). *Acta Cryst.* D**73**, 1020–1028.

Zheng, M., Reimers, J. R., Waller, M. P. & Afonine, P. V. (2017). *Acta Cryst.* D**73**, 45–52.