Hydrogen atoms constitute about half of all atoms in proteins and play a critical role in enzyme mechanisms and macromolecular and solvent structure. Hydrogen atom positions can readily be determined by neutron diffraction, and as such, neutron diffraction is an invaluable tool for elucidating molecular mechanisms. Joint refinement of neutron and X-ray diffraction data can lead to improved models compared with the use of neutron data alone and has now been incorporated into modern, maximum-likelihood based crystallographic refinement programs like CNS. Joint refinement has been applied to neutron and X-ray diffraction data collected on crystals of diisopropyl fluorophosphatase (DFPase), a calcium-dependent phosphotriesterase capable of detoxifying organophosphorus nerve agents. Neutron omit maps reveal a number of important features pertaining to the mechanism of DFPase. Solvent molecule W33, coordinating the catalytic calcium, is a water molecule in a strained coordination environment, and not a hydroxide. The smallest Ca–O–H angle is 53°, well beyond the smallest angles previously observed. Residue Asp-229, is deprotonated, supporting a mechanism involving nucleophilic attack by Asp-229, and excluding water activation by the catalytic calcium. The extended network of hydrogen bonding interactions in the central water filled tunnel of DFPase is revealed, showing that internal solvent molecules form an important, integrated part of the overall structure.

As a test of the utility of this development, we used nCNS to refine the structure of diisopropyl fluorophosphatase (DFPase) from the squid Loligo vulgaris (EC 3.1.8.2) using both X-ray and neutron crystallographic data. DFPase, a 35 kDa, calcium-dependent phosphotriesterase, efficiently hydrolyzes a wide variety of highly toxic organophosphorus compounds, although its physiological substrate remains unknown (16, 17). In addition to hydrolyzing diisopropyl fluorophosphate (DFP), the enzyme detoxifies a range of organophosphorus nerve agents, including Tabun, Sarin, Soman, and Cyclohexylsarin, through hydrolysis of the bond between the phosphorus and the leaving group (17). DFPase can be highly expressed, is thermally stable, and as such, can be precisely oriented in nuclear density maps, whereas hydrogen atoms in water molecules are usually invisible in electron density maps. This allows the unambiguous assignment of hydrogen bonding interactions.

Nevertheless, the practical application of neutrons to protein crystallography has been limited in the past by the relatively weak neutron flux of beam lines compared with synchrotrons for X-ray experiments and the requirement for large crystals (>1 mm³). The low neutron flux leads to extended data acquisition times that in addition to the very small number of available sources limits the availability of beam time for individual experiments. A further challenge is the refinement of the structural model against neutron data, due to its limited resolution and the low data-to-parameters ratio. The data-to-parameters ratio is particularly critical, because the addition of hydrogen atoms in a macromolecule roughly doubles the number of atoms whose position and temperature factor are being refined. Inclusion of both X-ray and neutron data in refinement may aid in increasing the effective number of observations. A series of studies ~25 years ago demonstrated the potential of the joint X-ray and neutron refinement technique. The implementation of joint X-ray and neutron refinement into programs such as CNS (12) [in the form of a patch version called nCNS (13)] and PHENIX (14, 15) has now made it possible to combine this approach with recent developments in macromolecular structure refinement, such as cross-validated maximum likelihood and simulated annealing.

As a test of the utility of this development, we used nCNS to refine the structure of diisopropyl fluorophosphatase (DFPase) from the squid Loligo vulgaris (EC 3.1.8.2) using both X-ray and neutron crystallographic data. DFPase, a 35 kDa, calcium-dependent phosphotriesterase, efficiently hydrolyzes a wide variety of highly toxic organophosphorus compounds, although its physiological substrate remains unknown (16, 17). In addition to hydrolyzing diisopropyl fluorophosphate (DFP), the enzyme detoxifies a range of organophosphorus nerve agents, including Tabun, Sarin, Soman, and Cyclohexylsarin, through hydrolysis of the bond between the phosphorus and the leaving group (17). DFPase can be highly expressed, is thermally stable, and as such, can be precisely oriented in nuclear density maps, whereas hydrogen atoms in water molecules are usually invisible in electron density maps. This allows the unambiguous assignment of hydrogen bonding interactions.

Nevertheless, the practical application of neutrons to protein crystallography has been limited in the past by the relatively weak neutron flux of beam lines compared with synchrotrons for X-ray experiments and the requirement for large crystals (>1 mm³). The low neutron flux leads to extended data acquisition times that in addition to the very small number of available sources limits the availability of beam time for individual experiments. A further challenge is the refinement of the structural model against neutron data, due to its limited resolution and the low data-to-parameters ratio. The data-to-parameters ratio is particularly critical, because the addition of hydrogen atoms in a macromolecule roughly doubles the number of atoms whose position and temperature factor are being refined. Inclusion of both X-ray and neutron data in refinement may aid in increasing the effective number of observations. A series of studies ~25 years ago demonstrated the potential of the joint X-ray and neutron refinement technique. The implementation of joint X-ray and neutron refinement into programs such as CNS (12) [in the form of a patch version called nCNS (13)] and PHENIX (14, 15) has now made it possible to combine this approach with recent developments in macromolecular structure refinement, such as cross-validated maximum likelihood and simulated annealing.

As a test of the utility of this development, we used nCNS to refine the structure of diisopropyl fluorophosphatase (DFPase) from the squid Loligo vulgaris (EC 3.1.8.2) using both X-ray and neutron crystallographic data. DFPase, a 35 kDa, calcium-dependent phosphotriesterase, efficiently hydrolyzes a wide variety of highly toxic organophosphorus compounds, although its physiological substrate remains unknown (16, 17). In addition to hydrolyzing diisopropyl fluorophosphate (DFP), the enzyme detoxifies a range of organophosphorus nerve agents, including Tabun, Sarin, Soman, and Cyclohexylsarin, through hydrolysis of the bond between the phosphorus and the leaving group (17). DFPase can be highly expressed, is thermally stable, and as such, can be precisely oriented in nuclear density maps, whereas hydrogen atoms in water molecules are usually invisible in electron density maps. This allows the unambiguous assignment of hydrogen bonding interactions.

Nevertheless, the practical application of neutrons to protein crystallography has been limited in the past by the relatively weak neutron flux of beam lines compared with synchrotrons for X-ray experiments and the requirement for large crystals (>1 mm³). The low neutron flux leads to extended data acquisition times that in addition to the very small number of available sources limits the availability of beam time for individual experiments. A further challenge is the refinement of the structural model against neutron data, due to its limited resolution and the low data-to-parameters ratio. The data-to-parameters ratio is particularly critical, because the addition of hydrogen atoms in a macromolecule roughly doubles the number of atoms whose position and temperature factor are being refined. Inclusion of both X-ray and neutron data in refinement may aid in increasing the effective number of observations. A series of studies ~25 years ago demonstrated the potential of the joint X-ray and neutron refinement technique. The implementation of joint X-ray and neutron refinement into programs such as CNS (12) [in the form of a patch version called nCNS (13)] and PHENIX (14, 15) has now made it possible to combine this approach with recent developments in macromolecular structure refinement, such as cross-validated maximum likelihood and simulated annealing.

As a test of the utility of this development, we used nCNS to refine the structure of diisopropyl fluorophosphatase (DFPase) from the squid Loligo vulgaris (EC 3.1.8.2) using both X-ray and neutron crystallographic data. DFPase, a 35 kDa, calcium-dependent phosphotriesterase, efficiently hydrolyzes a wide variety of highly toxic organophosphorus compounds, although its physiological substrate remains unknown (16, 17). In addition to hydrolyzing diisopropyl fluorophosphate (DFP), the enzyme detoxifies a range of organophosphorus nerve agents, including Tabun, Sarin, Soman, and Cyclohexylsarin, through hydrolysis of the bond between the phosphorus and the leaving group (17). DFPase can be highly expressed, is thermally stable, and as such,
A second possible mechanism involves activation of calcium-bound water by the metal, resulting in a hydroxide nucleophile that can then attack the substrate (Fig. 2B). However, because the active-site metal in DFPase can be exchanged for other divalent metal ions such as magnesium and barium, with little change in catalytic activity, this mechanism has been deemed to be highly unlikely (19). Nevertheless, a similar mechanism was proposed for the lactonase activity of the structurally homologous enzyme Drp35 (20). Here, the calcium-bound water is activated by one of the metal-coordinating aspartate residues resulting in a hydroxide nucleophile.

To discriminate between these proposals, it is necessary to accurately determine the protonation states of the active-site residues, and the orientations and identities of critical solvent molecules in the active site vicinity. In addition to assigning the protonation states of residues Asp-229 and Glu-21 in the active site, a particularly important question with a direct bearing on mechanism is whether the solvent molecule W33, bound to the catalytic calcium, is a water molecule or a hydroxide ion. For our proposed mechanism, we expect Asp-229 to be deprotonated and W33 to be a water molecule, whereas in the case of metal-assisted water activation, we expect W33 to be a hydroxide ion. Although the X-ray structure of the holoenzyme has been proposed for the lactonase activity of the structurally homologous enzyme Drp35 (20), Here, the calcium-bound water is activated by one of the metal-coordinating aspartate residues resulting in a hydroxide nucleophile.

We present here the results of a neutron protein crystallography study of DFPase. In this work, we have successfully collected time-of-flight Laue neutron data to 2.2 Å from a small (0.43 mm³) crystal equilibrated against D₂O-containing mother liquor, using the PCS beam line (26) at the Los Alamos Neutron Science Center (LANSCE) spallation source. Because H atoms account for nearly half of the atoms in a protein, adding H atoms in a neutron structure refinement increases the number of parameters and reduces the data-to-parameter ratio, increasing the danger of overfitting and decreasing the accuracy of the optimized model. We have therefore used the approach, originally developed by Wlodawer and Hendrickson (27), of combining the neutron data with 1.8-Å X-ray data in a joint (XN) structure refinement. We demonstrate the effectiveness of the joint refinement in nCNS at determining the protonation states of residues and solvent molecules in the active site of DFPase. We show the orientation of water molecules and the associated hydrogen bonding network in the water tunnel of DFPase.

We discuss the implications of our results for the phosphotriesterase mechanism of DFPase, including a comparison of the competing potential reaction mechanisms for this class of enzymes. As we proposed earlier, both DFPase and PON may share the same phosphotriesterase mechanism in light of the highly conserved coordination environment around the catalytic calcium. Therefore, mechanistic insights derived from the XN structure of DFPase may aid in the understanding of the phosphotriesterase activity of PON.
Results and Discussion

Joint Neutron and X-Ray Refinement (XN). X-ray and neutron crystallographic data provide complementary structural information. Labile hydrogen atoms, including those on solvent molecules, hydroxyl groups, acidic hydrogens that might be absent due to deprotonation, and other exchangeable hydrogens that might be substituted by deuterium, are accurately located through neutron diffraction data. These hydrogens are often invisible in electron density maps. Conversely, X-ray data can also help define heavy atom positions and a limited number of hydrogens at high resolution. For example, methylene groups are rarely visible in neutron maps, because the opposite signs of the scattering lengths of C and H cancel each other out (Fig. 3A). In this case X-ray data not only defines the position of the carbon atom, but from inclusion of prior knowledge about bond lengths and angles in the refinement also helps to define the hydrogen positions. Joint refinement can thus lead to improved phases and better electron and nuclear density maps. Wlodawer and Hendrickson observed a quarter-century ago during their refinement of ribonuclease A, that refinement of neutron data on its own was often unstable, and that including the X-ray data in a joint refinement greatly reduced errors in the model (27).

Despite the advantages of joint X-ray and neutron refinement, to date most neutron structures have been determined by using programs, such as SHELX (28) and CNS (12), which allow the use of either X-ray or neutron data, but not both at the same time. In addition, these programs were originally developed for X-ray crystallography and are not completely suitable for macromolecular neutron crystallography. Joint X-ray and neutron refinement for biological macromolecules has now been implemented in a practical setting, through the development of nCNS (13), which allows for a robust, user-friendly crystallographic refinement environment.

XN Structure of DFPase. Neutron diffraction on a crystal of DFPase, exchanged against deuterated mother liquor by vapor diffusion, allowed for the determination of the complete structure of the enzyme, including H atoms and the ordered solvent network. Compared with the X-ray structure, the model has approximately double the number of atoms, due to the inclusion of H atoms. To account for the larger number of parameters, the refinement of the structure used a patched version of CNS (nCNS) that allows for refinement of the model against both X-ray and neutron diffraction data. As such, the inclusion of X-ray and neutron data increased the data-to-parameters ratio and improved the accuracy of the model, while reducing the possibility of overfitting of data. Details of the joint refinement are described in SI Text, and data collection and refinement statistics are summarized in Table S1.

Inspection of the neutron maps, at a resolution of 2.2 Å, showed clear features for nuclear density. Because of the cancellation of the scattering of H and N atoms, unexchanged backbone amides show a lack of positive nuclear density (Fig. 3A). Water molecules, in the form of D2O, appear as bean-shaped density (Fig. 3B). The protonation state of histidine residues can be seen (Fig. 3C), whereas ND2 and ND3 groups at the termini of side chains appear as strong nuclear density (Fig. 3B).

The degree of backbone amide exchange was determined by individual occupancy refinement of main-chain amide hydrogens (Fig. 1). Labile hydrogens in the crystal were exchanged for deuteriums for a period of less than 1 month before the start of data collection. The observed pattern of exchange mirrors the results obtained from NMR studies, which showed that ~30% of the backbone amides did not show any detectable exchange (29). Although internal portions of the protein are generally most resistant to exchange (30), surprisingly, a number of surface residues do not undergo exchange. These residues correspond to β-strand amide hydrogens involved in hydrogen bonding with adjacent β-strands, whereas amide protons pointing toward the solvent are exchangeable.

The final model is essentially identical to the room temperature X-ray structure (PDB: 2GVW), with a root mean squared deviation (rmsd) of 0.086 Å over Ca atoms. The differences between the joint refined structure and 2GVW lie primarily in...
residues with larger side chains; for example, the presence of strong nuclear density for the terminal ND$_3^+$ groups on lysines allows for unambiguous placement of such side chains. There are greater differences between the joint refined structure and the ultrahigh resolution structure (Protein Data Bank ID code: 1PJX), with a rmsd of 0.258 Å over the Ca atoms, attributable to the different temperature of data collection (100 K vs. room temperature) and slightly different cell dimensions.

DFPase is the first protein for which neutron diffraction data are reported that contains an extended network of internal water molecules, connected by hydrogen bonds in form of a water tunnel in the center of the propeller. Despite the highly ordered water orientations within the tunnel and the series of "gates" in the tunnel such as the 2 calcium ions and the Asp-121/Asn-272 water orientations within the tunnel and the series of "gates" in the tunnel such as the 2 calcium ions and the Asp-121/Asn-272 hydrogen bond, all waters in the tunnel have exchanged for D$_2$O within the time scale of the experiment. The tunnel segment between structural calcium Ca2 and catalytic calcium Ca1 shows a complex network of interactions (Fig. 4), involving a variety of side chain–water and main chain–water hydrogen bonds (Fig. S1). The active site end of the water tunnel is occupied by catalytic calcium Ca1. Water molecules W34 and W32, coordinating with Ca1 on the inside of the tunnel, are unusual because they interact only with Ca1 and the protein atoms on the sides of the tunnel wall and can therefore be considered as trapped or at least isolated. W34 donates a hydrogen bond to O$_{\text{am}}$ of Asp-121. The carboxylate group of Asp-121 is the starting point of an extended and intricate network of hydrogen bonded water molecules that passes through the protein and incorporates structural calcium Ca2, before forking into 2 exit channels. Asp-121 is connected to Ca2 through double water bridges that form hexagonal and pentagonal faces that share common edges, reminiscent of a clathrate structure. Perhaps significantly, all of the water molecules in this motif have hydrogen bond interactions only with main chain amide and carbonyl groups, and therefore this motif may be resistant to some mutations. This in agreement with the limited degree of sequence identity among β-propeller proteins (31).

**Active Site Environment and Mechanistic Implications.** The DFPase active site is an elongated cleft with the catalytic calcium Ca1 in the center. The active site is oriented such that 2 acidic amino acids, Asp-229 and Glu-21, coordinate the calcium on one side of the cleft, whereas Asn-120 and Asn-175 are oriented so that their side chain O$_{\text{am}}$ atoms coordinate to Ca1 (Fig. S4). Above the catalytic calcium is W33. Hydrogen bonds from the amino groups of Asn-120 and Asn-175 to W33 are possible in terms of distance and at least for Asn-120 in terms of geometry, although such interactions between a metal coordinated amide and a metal coordinated water molecule have rarely been reported (32).

O$_{\text{am}}$ of Asp-229 and O$_{\text{am}}$ of Glu-21 coordinate with Ca1. O$_{\text{am}}$ of Glu-21 accepts H bonds from the amino group of Asn-272 and the amide D of Gly-22, and O$_{\text{am}}$ of Asp-229 accepts a hydrogen bond from the hydroxyl D of Ser-271 and from W140. The nuclear density clearly shows that Asp-229 is deprotonated, which is further confirmed by the hydrogen bonding network in the active site region. Again, the joint refinement clearly reveals the orientation of the side chains, the fact that Asp-229 and Glu-21 are charged, and the distribution of protons in the hydrogen bonding arrangement. At the bottom of the active site pocket W32 and W34 coordinate with Ca1 and donate hydrogen bonds to amino acid side chains.

In enzymes, metal ions can serve as water activators in general base-catalyzed mechanisms. Solvent molecule W33 coordinates
the catalytic calcium, raising the question as to whether it is water or hydroxide. When interpreting $2F_{o} - F_{c}$ nuclear density during model building, W33 was clearly better modeled as a water molecule and not a hydroxyl ion (Fig. 5A). To unambiguously identify W33, a simulated annealing omit map was calculated, omitting the 2 deuteriums of W33. This omit map shows 2 difference $F_{o} - F_{c}$ peaks of equal intensity appearing in the geometry of a water molecule, further evidence that W33 is a water, and not a hydroxide (Fig. 5B). The coordination of W33 is unexpected, with the hydrogen atom in close proximity to Ca1. The observed distance between Ca1 and the O of W33 is 2.52 Å, whereas the closest hydrogen atom of W33 is only 2.08 Å from Ca1 (Fig. 5B). This unfavorable orientation for the hydrogen atom appears to be a consequence of the environment around W33, such that even more unfavorable orientations are minimized. Einspahr and Bugg (33) concluded that the minimal limit for Ca–O–H angles in crystalline hydrates is 75°, based on neutron and X-ray diffraction of calcium salts. In our case we observe an angle of 53° (Fig. 5B). We attribute this to the special environment in the protein active site, which is not comparable to highly ordered small inorganic molecule complexes. Although what we observe is unusual, and to our knowledge is the first such example reported for a calcium-bound water molecule both in proteins and small molecules, the omit map $F_{o} - F_{c}$ density is clear and unambiguous. It is important to note that in the present orientation of W33 there is no hydrogen bonding interaction with the O$^{61}$ of Asp-229, which could enhance water activation.

Asp-229 is deprotonated in the structure as required for a nucleophilic attack on a substrate molecule, consistent with our earlier isotope labeling studies. Finally, this observed orientation would not have been seen with even the highest resolution X-ray structures available, but is unambiguous with neutron diffraction at moderate resolution.

Furthermore, there are a number of supporting chemical arguments why W33 is water, and not hydroxide. It was reported in the literature that in the case of magnesium complexes water is indeed activated by coordination to the metal; however, the presence of even a single carboxylic ligand in the complex increases the free energy of the deprotonation step by ~80 kcal/mol (34). In the case of DFPase there are 2 carboxylates in the coordination sphere of the catalytic calcium, further increasing the free energy of deprotonation. The interpretation of W33 as a water molecule is also consistent with the finding that little activation by the metal ion in PON. At present neutron diffraction experiments on PON do not seem feasible given the small crystal size reported for the solved PON structure and taking into account unit cell parameters and space group (21). The XN structure of DFPase might therefore serve as a valuable model and inspire further studies to establish the phosphotriesterase mechanism for PON.

Implications for Neutron Protein Crystallography. Neutron structures have been reported for ~20 individual proteins and nucleic acids to date, compared with 52,000 X-ray structures deposited in the Protein Data Bank as of the end of July 2008. The XN structure of DFPase is one of the first examples of joint X-ray and neutron structure refinement for a protein in combination with cross-validated maximum likelihood simulated annealing. The use of this approach has several notable advantages, including an increased data-to-parameters ratio and more accurate information about protein structure, than either X-ray or neutron crystallography on its own. However, we found that one of the clearest practical benefits of joint refinement in this particular case was during water building. Superposition of the electron and nuclear density maps clearly indicated the orientation of water molecules, whereas inspection of either map on its own was often ambiguous. Because X-ray diffraction data normally extend to higher resolution than neutron data, the refinement can potentially be biased toward the positions of the heavy atoms. This illustrates the importance of proper weighting factors for X-ray and neutron data during refinement. Wlodawer and Hendrickson observed that refinement of the neutron data on its own was sometimes unstable (27); thus the addition of the X-ray data helped 'stabilize' the structure, and the neutron data allowed for the refinement of the positions of the hydrogen atoms.

Currently, the number of neutron sources is a limiting factor, and there has been a significant increase in the number of people interested in screening crystals for neutron diffraction. Sample sizes have become smaller, and these factors will inevitably lead to difficult and weak data. The use of nCNS provides a convenient platform for joint refinement, and employs the robust maximum likelihood target function, ideally suited for datasets with relatively weak data, which is typically the case for neutron data collected with the time-of-flight technique at spallation sources. The useful features of omit maps calculated in nCNS are shown, which allowed for the determining the orientation of W33. Furthermore, we note that most neutron structures solved since 2007 have applied the joint refinement technique, and thus will be the choice method in the future for refinement of neutron diffraction data. The implementation we discuss here is thus a major step forward, providing a starting point for the development of new strategies to further improve the refinement of neutron structures. The Protein Data Bank is in the process of determining a standard format for the deposition of future joint refined structures.

In addition to the insights into the DFPase mechanism reported here, this work also represents a major technical advancement in neutron crystallography. The crystal used here was 0.43 mm$^3$ in size, among the smallest sample size used so far for neutron crystallography (36). The crystal was exchanged for less than 1 month before the start of data acquisition. In light of these improvements in sample size requirements, data acquisition
modified for time-of-flight data (26). The structure was refined by using a version of CNS (12), called nCNS, modified for joint X-ray and neutron refinement (13). The refinement used neutron data between 20.0–2.2 Å resolution, accounting for a slight difference in the number of reflections than reported (36). A detailed treatment of the joint refinement method is described in the SI Text, and statistics of the joint refinement of DFPase are reported in Table S1.

ACKNOWLEDGMENTS. We thank Dr. Leighton Coates for expert advice on refinement and Mary Jo Waltman for technical assistance. We thank Drs. Jenny Glusker and Alexander Koglin for helpful discussions. This project was funded by the Hessisches Ministerium für Wissenschaft und Kultur and the German Ministry of Defense under Grants EES90620044F70 and EUR3G/6G15S/EB081. A travel grant was provided to M.-M.B. and J.C.-H.C. by the Glaxo Smith-Kline Foundation. The PCS is funded by the Office of Science and the Office of Biological and Environmental Research of the U.S. Department of Energy. M.M. and P.L. were partly supported by an National Institutes of Health (NIH) National Institute of General Medical Sciences-funded consortium (1R01GM071939–01) between Los Alamos National Laboratory and Lawrence Berkeley National Laboratory to develop computational tools for neutron protein crystallography.


Materials and Methods

Data Collection, Processing, and Refinement. Room-temperature neutron data were collected and processed as reported (36). Briefly, a single crystal measuring 2.4 × 0.5 × 0.36 mm in size was mounted in a glass capillary, and labile hydrogen atoms served as detectors for deuteration by vapor diffusion for 1 week before the start of data collection. Thirty-seven time-of-flight neutron diffraction images were collected at the PCS at LANSCE, with exposures of up to 24 h per image. Data were processed by using a version of d*TREK (37)