

Spectroscopic and inclusion properties of G-series chemical warfare agents and their simulants: A DFT study

Mark R. Sambrook,^a Ian A. Gass^b and Peter J. Cragg^{b*}

^a*CBR Division, Dstl Porton Down, Salisbury, SP4 0JQ, UK*

^b*School of Pharmacy and Biomolecular Sciences, Huxley Building, University of Brighton, Brighton BN2 4GJ, UK*

Contact Peter J. Cragg *P.J. Cragg@brighton.ac.uk*

Mark R. Sambrook orcid.org/0000-0002-4826-860X

Ian A. Gass orcid.org/0000-0001-8951-7847

Peter J. Cragg orcid.org/0000-0002-6134-2093

Spectroscopic and inclusion properties of G-series chemical warfare agents and their simulants: A DFT study

A computational protocol to predict the infrared spectra of chemical warfare agents (CWAs) tabun (GA), sarin (GB), soman (GD) and cyclosarin (GF) has been developed. Sarin was used to benchmark the method through gas phase simulations. DFT calculations using the EDF2 functional and diffuse 6-311++G** basis set was found to give the closest match to experimental infrared spectra. Using the same functional the 6-31G (2df, 2p) basis set was found to be superior when hydrated sarin was modelled. GA, GB, GD and GF, together with 11 commonly used simulants, were modelled in the gas and hydrated states. Complexes of GB and a number of CWA mimics with α -cyclodextrin were modelled to give insight into their different modes of inclusion.

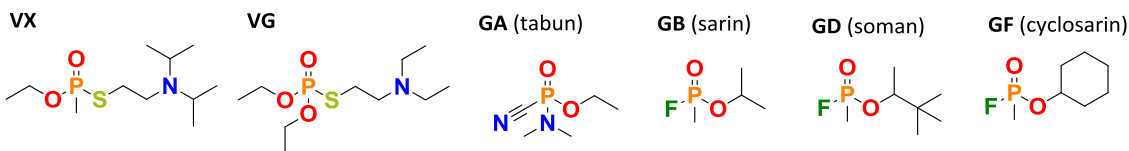
Keywords: density functional theory, chemical warfare agents, cyclodextrin, infrared spectra, inclusion complex

Introduction

The history of the development and use of Chemical Warfare Agents (CWAs) through, predominantly, the early and mid-twentieth century is well documented, and in 1993 the Chemical Weapons Convention banned the manufacture of such materials (1). Nevertheless, they have gained prominent attention over recent decades due to evidence of their use in Iraq in 1987 (2), Japan in 1994 (3) and 1995 (4), and Syria in 2013 (5). In the event of deliberate or accidental release of CWAs it is essential to detect and destroy these highly toxic materials. A large number of hydrolytic, enzymatic and thermal destruction/degradation methods have been proposed (6-8) as have detection systems based on spectroscopic or spectrometric approaches (9, 10). However, recent interest in the supramolecular chemistry of CWAs has led to a significant number of reports exploring the molecular recognition properties of these molecules by functional materials. Such materials can act as sensors, destructive catalysts and responsive gels (11-15). Chemical agent simulants, with related structural or physical properties but with much lower toxicity than the CWAs, are often used by research groups interested in developing novel detection technologies or functional materials that may find

application in future protection or decontamination systems. Increased reliance on the use of simulants to develop promising technologies requires an improved understanding of how the properties of CWA and simulants correlate.

CWAs



Simulants

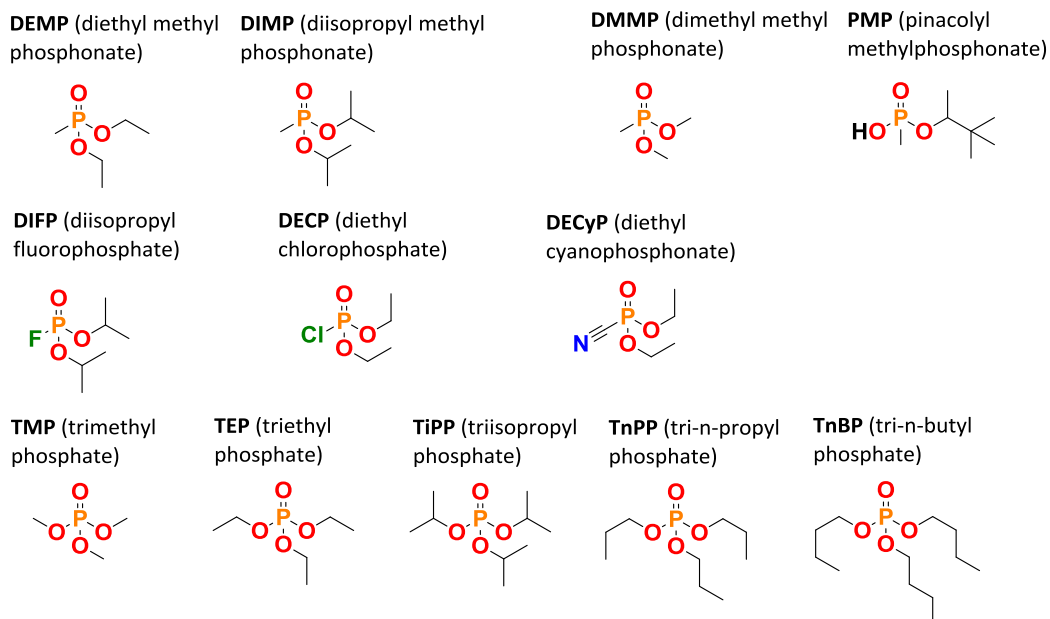


Figure 1. Chemical warfare agents and simulants

Key to developing novel methods of sensing and destroying CWAs through the formation of inclusion complexes is the ability of researchers to test prospective methods on simulants with the knowledge that they accurately reflect the physicochemical behaviour of CWAs most relevant to the task. Alternatively, there should be a clear understanding of how simulant data could be more effectively extrapolated to predict CWA behaviour; rare examples of CWA-simulant comparisons include the coordination of such molecules to lanthanides (11, 12), and the triggering of stimuli-responsive gels (15).

To more closely match the molecular properties of simulants to CWAs it is essential that the salient features of each CWA are also present in its selected simulant. At the most basic level these will include molecular structure and geometry. Reactivity of the phosphorus bonds, together with the chirality of the phosphorus atom, underlies the toxicity of CWAs. Whilst this reactivity should not be reproduced in a simulant, having similar vibrational modes should make it interact with sensors or decontaminants in a similar way. For example, if simulant and CWA P=O groups have similar vibrational frequencies then their ideal P-O bond lengths, electronic polarisation and predilection to interact with hydrogen bond donors will be almost identical. Consequently, when subjected to the same environment, such as inclusion within macrocycles or binding to imprinted polymers, simulants will give the same response as the CWA. The first step is therefore to develop a computational protocol that accurately predicts the geometries and vibrational modes of both CWAs and simulants. Once this has been achieved, the same approach can be used to investigate the reasons for the differences between CWA and simulant binding to substrates. Thus, these models can be used to predict the inclusion complexes which CWAs can form and, ideally, predict which simulants will form similar complexes. From this it will be possible to use appropriate simulants in place of CWAs with assurance that their inclusion behaviour will be similar when binding to receptors that form part of a sensor.

Simple calculations to determine the gas phase structures of CWAs and simulants have been reported since 1985 when Ewig and Van Wazer modelled the *ab initio* structure of *O*-methyl methylphosphonofluoridate with the STO-3G and 321-G* basis sets (16). Subsequent work focused on either calculation of microwave spectra (17-19), of use when designing microwave-based sensors to detect CWAs, or surface interactions relevant to remediation or transport through porous materials (20-27).

Computer simulations generally consider gas phase structures. While these are of interest, particularly for sensor applications, it is also important to understand what occurs in aqueous solution, as investigated by the groups of Ault and Neimark, because atmospheric moisture is likely to affect detection (28-30).

There has been limited investigation of simulations to replicate infrared spectra of CWAs and simulants. Mott considered V- and G-series compounds as well as simulants but all were modelled in the gas phase and show variable reproducibility of the P=O stretches in fluorinated derivatives (31). Cuisset modelled simulants DMMP, TMP and TEP at MP2/6-311++G(d,p) and DFT/B3LYP/6-311G++(3df,2p) levels of theory to extract infrared data and attempt to discriminate between species and conformers (32). Alam and Pearce (33) investigated the effects of explicit microsolvation of GB using DFT/B3LYP/6-311G++(2d,2p) to determine if the level of solvation had a significant effect on the computed infrared spectra. Although the inclusion of water in the simulation improved the fit of the P=O stretch to experimental data little difference was seen between simulations of one, two, three or four solvent molecules. Zhang et al. (34) used DFT methods to model agents VX, VG, GA, GB and GD together with simulants DMMP and dimethyl (3,3-dichloroallyl)phosphonate (DDV). B3LYP/6-311+G(d,p) gave the best prediction of infrared spectra but underestimated frequencies associated with bonds to phosphorus. Hameka et al. (35, 36) applied the HF/6-31G* method to model methylphosphonates and related species but required further correction factors to match experimental infrared spectra. A higher level (MP2/aug-cc-pVTZ) study of GB, undertaken to investigate its thermal decomposition mechanism, gave good agreement with experimental data from 750 and 1500 cm^{-1} but still overestimated C-H stretching modes by *ca.* 100 cm^{-1} (37).

Herein, we report on a combination of semi-empirical and DFT approaches to model CWAs and their simulants with the aim of generating structures that accurately reproduce experimentally determined infrared spectra. These models can be used to give insight into complexation of the different species and reasons why simulants often fail to bind in the same manner as the agents. To determine if the formation of supramolecules incorporating CWAs and simulants involved similar interactions, α -cyclodextrin (α -CD) inclusion complexes of GB and selected simulants were modelled.

Methods

Calculations were undertaken using the *Spartan '16 Parallel Suite (38)* running on a Mac Pro with 3.5 GHz 6-Core Intel Xenon E5 processors and two threads per core. CWAs and simulants were constructed using the Build option. A full conformational search of the CWA structure using molecular mechanics (MMFF) followed by geometry refinement by DFT (EDF2/6-31G then 6-311++G**). Vibrational frequencies were calculated after the final DFT geometry refinement. For solvated structures the same method was applied but with the addition of 55 water molecules to simulate a molar aqueous solution. Following a semi-empirical geometry refinement (PM6), all non-hydrogen-bonding water molecules were removed and the first DFT calculation (EDF2/6-31G) undertaken with the oxygen atoms of the associated water molecules frozen. After the second DFT step (EDF2/6-311++G**) the water molecules were deleted and the single point energy calculated for the CWA structure. This geometry was used to derive vibrational frequencies which incorporate the effects of hydrogen bonding to the CWA. The α -CD host was geometry optimised by semi-empirical methods from its X-ray coordinates (39) before refinement by HF/3-21G then DFT/EDF2/6-311G.

Results and Discussion

Validation of Molecular Structure

GB in the Gas and Liquid Phase

The optimised geometry of GB was calculated at increasing levels of computational sophistication. The MMFF forcefield was used for the molecular mechanics simulation, PM6 was used for semi-empirical calculations and a range of functionals and basis sets were used for DFT calculations. The geometry optimised structures from each of these simulations were used to generate infrared spectra (see Supplementary Information, S5-S13). Two factors were considered: was the simulated spectrum in broad agreement with the experimental spectra and did the peaks due to the P=O stretch and asymmetric P-O-R stretch correlate with the experimental values? Experimental gas phase spectra from Piffath (40), Sharpe (41), Heiss (42) and liquid phase Dstl data (see Supplementary Information, S3) were used for reference.

From inspection of the data presented in Table 1 it is clear that peaks due to C-H vibrations in the region just below 3000 cm^{-1} are poorly replicated, except by MMFF, which is an acknowledged issue with infrared simulations (31). Consequently the fingerprint region is of greater analytical importance. The best results, which most closely matched data from Piffath (40), were obtained using the DFT/EDF2/6-311++G** model (Figure 2).

Gas Phase Calculations for CWAs and Simulants

Following on from the method adopted to model GB, three other chemical warfare agents (GA, GD and GF) and 11 simulants (DEMP, DIMP, DMMP, DIFP, DECP, DECyP, TMP, TEP, TiPP, TnPP and TnBP) were modelled in the gas phase (Tables 2, 3 and 4). Of these the calculated infrared spectra for the G-series were in good

agreement with experimental values. The simulants, other than DMMP and DEMP, also modelled well.

	Expt (Dstl)	Expt^[40]	MMFF	PM6	EDF2/6-311++G**
IR (cm⁻¹)					
C-H str	2977 + 2930	2986 + 2935	2991 + 2989	2845 + 2799	3118 + 3043
CH₃ def	1497 + 1465	1469 + 1458	1463 + 1457	1204	1489 + 1475
P-CH₃ asym def	1418	1420	1449	1336	1419
CH₃ i-pr def	1385 + 1342	1389 + 1379	1438 + 1430	1343 + 1328	1398 + 1361
P-CH₃ sym def	1315	1321	1416	1200	1345
P=O str	*	1278	1254	1191	1294
P-O-i-pr	1136 + 1121	1180 + 1145 + 1106	1147 + 1137 + 1101	1162 + 1158	1199 + 1164 + 1128
P-O-C str	994	1015	1060	1038	1008
P-CH₃ rock	936	924 + 907	898 + 883	955	940 + 924
P-F str	874	838	930	756	809
P-O-C bend	810	778	743	-	763
P-C str	754	723	724	-	716

*experimental P=O stretch obscured by a shoulder from 1300-1124 cm⁻¹

Table 1. Calculated gas phase spectra vs experimental liquid phase spectra for GB.

Hydrated GB

The CWAs will always encounter water molecules in the natural environment, if only from humid air. To explicitly include the effect of water molecules on GB, simulations were constructed in which a single GB molecule was surrounded by 55 water molecules, approximating a molar solution, and geometry optimization undertaken using PM6 from a MMFF starting geometry. The resulting simulation was inspected and all water molecules not hydrogen bonded to GB were removed. Geometry optimization by DFT/EDF2/6-31G* followed and the output inspected again. Any water molecules no longer within hydrogen bonding distance were removed. The oxygen atoms of those remaining water molecules were frozen and optimization resumed using

DFT/EDF2 with either the 6-311++G** or 6-311G (2df, 2p) basis set. Following this final optimization step, the water molecules were removed and a single point energy calculation was used as the input to generate the infrared spectrum. Incorporating p-orbital derived diffuse effects in the latter basis set appeared to give a better correlation with the Piffath and Sharpe spectra when the hydrated model was adopted (see Table 5 and Supplementary Information, S10). Furthermore, the absorbance peak shifts for P-O-R and P=O bonds were consistent with those observed for the simulant DMMP when it was hydrated (40).

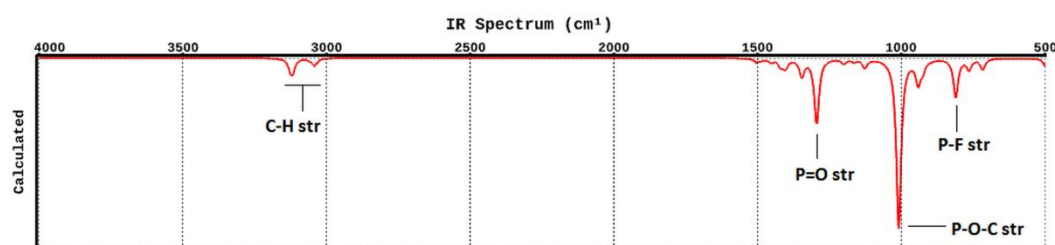
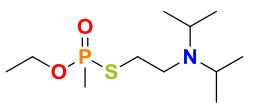
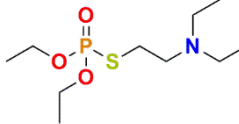
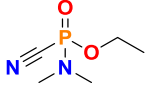
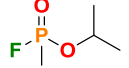
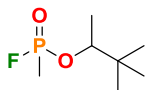
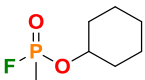


Figure 2. Simulated gas phase spectrum of GB (DFT/EDF2/6-311++G**)

CWA	Structure	IR expt ^[40, 42] (P=O, P-O-R)	IR calcd ^a (P=O, P-O-R)	IR calcd ^[31] (P=O, P-O-R)
VX		1227, 1031	1239, 1073	1232, 1067
VG		1255, 1012 + 966	1253, 1053 + 966	1259, 1063 + 1040
GA (tabun)		1269, 1030	1271, 1054	1267, 1054
GB (sarin)		1278, 1015	1294, 1008	1308, 1015
GD (soman)		1283, 1018	1291, 985	1304, 1025 + 987
GF (cyclosarin)		1281, 1058 + 1032	1299, 1038 + 998	1308, 1048 + 1026

^athis work, DFT/ EDF2/6-311++G**

Table 2. Gas phase data for CWAs.

Simulant	Structure	IR expt ^[40,41] (P=O, P-O-R)	IR calcd ^a (P=O, P-O-R)	IR calcd ^[31] (P=O, P-O-R)
DMMP		1245, 1058 + 1032	1295, 1109 + 1085	1271, 1078 + 1057
DEMP		1276, 1050	1254, 973 + 962	1265, 1074 + 1051
DIMP		1245, 1012 + 963	1249, 1015 + 983	1242, 1005 + 977
DIFP		1297 + 903	1368, 1069 + 1037	1313, 1027 + 997
DECP		1298, 1028	1275, 1041 + 960	N/D
DECyP		1304, 1027	1289, 1052 + 957	N/D

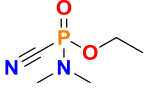
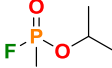
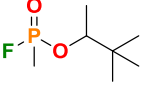
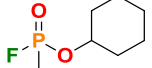
^athis work, DFT/ EDF2/6-311++G**, N/D not determined

Table 3. Gas phase data for dialkyl phosphonate and phosphate simulants

Simulant	Structure	IR expt ^[40] (P=O, P-O-R)	IR calcd ^a (P=O, P-O-R)
TMP		1292, 1060	1271, 1072
TEP		1280, 1050	1260, 1052 + 957
TiPP		1259, 980	1259, 999
TnPP		1268, 1044 + 986	1291, 1037 + 1020 + 1000
TnBP		1274, 1028 + 991	1291, 1044 + 1019

^athis work, DFT/ EDF2/6-311++G**

Table 4. Gas phase data for trialkyl phosphate simulants

CWA	Structure	IR calcd-hyd ^a (P=O, P-O-R)	IR expt ^[40] (P=O, P-O-R)	IR calcd ^[31] (P=O, P-O-R)
GA (tabun)		1252, 1036	1269, 1030	1267, 1054
GB (sarin)		1258, 1012	1278, 1015	1308, 1015
GD (soman)		1203, 957	1283, 1018	1304, 1025 + 987
GF (cyclosarin)		1290, 1013 + 1004	1281, 1058 + 1032	1308, 1048 + 1026

^athis work, DFT/EDF2/6-311G (2df, 2p)

Table 5. Hydrated CWA data.

Comparison with Other Simulations

The most relevant literature data come from Mott (31) who calculated infrared spectra for 10 CWAs and five simulants by DFT using the B3LYP functional and the 6-311+G (d,p) basis set. All calculations were in the gas phase and correlated well with experiment for the P-O-R asymmetric stretch, however, the P=O stretches were shifted to significantly lower frequencies for GB, GD and GF. Consideration of explicit microsolvation in the current study (Table 5) has given more accurate results, in agreement with data from Alam and Pearce for GB (33), and represents a better protocol for the G-series of CWAs. Indeed, it may be that microsolvation is a more appropriate method to model other CWAs and their simulants, however, due to the extended computational time required, it was not considered in this study.

Inclusion Complexes

The validated geometry optimisation methods for gas phase CWAs and simulants were applied to predict the formation of host-guest complexes of α -CD with GB and a range of simulants was

undertaken. X-Ray structures indicate two water molecules are generally bound inside the cavity (39) and, when modelled, the dihydrate inclusion complex proved stable. Water molecules were removed from the geometry optimised α -CD and the guest molecules introduced at the opening of the macrocycle's cavity. The *Minimize* (MMFF) routine within the *Spartan '16 Build* menu was used to generate an initial geometry optimization. Complexes were then modelled (HF/3-21G) and then refined (DFT/EDF2/6-311G*) to give a final geometry. Complexes were analysed to determine the orientation of the guest molecule and to what extent it was drawn into the host macrocycle cavity. Figures 3 to 6 and Tables 6 and 7 show how GB and its proposed simulants are bound by α -CD.

GB Inclusion in α -CD

While it is possible to use gas phase simulations to describe supramolecular complexes in organic solvents (43, 44), due to the extensive hydrogen bonding that could exist in the α -CD complexes an explicitly solvated approach was adopted. The simulation shows that the methyl group of GB is drawn into α -CD, leaving the isopropyl group outside the macrocycle's cavity, and the molecule held in place through hydrogen bonding to the phosphonate oxygen by two α -CD hydroxyl groups. These bifurcated O-H \cdots O interactions have an average H \cdots O distance of 1.94 Å (Table 6) and are similar in magnitude to both modelled (45) and experimental intramolecular α -CD hydrogen bonds (46, 47). This motif is strengthened by two further intramolecular hydrogen bonds to the two hydroxyl groups involved. The model supports proposed mechanisms by which α -CD mimics the action of acetylcholine esterase (AChE) in which initial binding of the phosphate oxygen to a hydroxyl group is followed by phosphorylation and finally hydrolysis to liberate isopropyl methylphosphonic acid (48, 49, 50). Désiré and Saint-André reported a dissociation constant K_d of 40 mM at pH 8.0 for the complexation of racemic GB by α -CD (51) so the simulations should predict successful complexation between GB and α -CD.

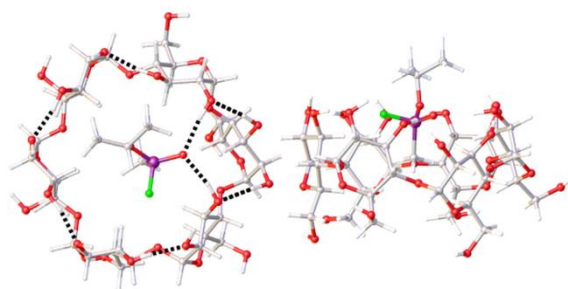


Figure 3. Top and side view of the α -CD complex with GB.

Complex	O-H...O		
	O...O(Å)	H...O(Å)	O-H...O(°)
GB to α-CD			
O-H...O	2.949	1.981	169.808
O-H...O	2.860	1.889	169.908
α-CD intramolecular			
O-H...O	2.759	1.792	167.615
O-H...O	2.815	1.845	167.986
O-H...O	2.805	1.833	169.231
O-H...O	2.815	1.844	170.122
O-H...O	2.874	1.915	166.302
O-H...O	2.809	1.859	162.550
DIMP to α-CD			
O-H...O	2.772	1.811	167.527
O-H...O	2.799	1.827	173.852
α-CD intramolecular			
O-H...O	2.851	1.878	170.502
O-H...O	2.684	1.792	149.072
O-H...O	2.776	1.829	162.613
O-H...O	2.881	1.921	166.033
O-H...O	2.841	1.874	168.219
O-H...O	2.761	1.801	166.122
DIFP to α-CD			
O-H...O	2.784	1.830	164.124
O-H...O	2.748	1.809	159.290
α-CD intramolecular			
O-H...O	2.798	1.819	170.427
O-H...O	2.717	1.751	165.701
O-H...O	2.790	1.841	162.875
O-H...O	2.924	1.957	169.992
O-H...O	2.853	1.889	166.502
O-H...O	2.749	1.765	175.343

Table 6. Intra- and intermolecular distances (Å) and angles (°) found in simulated α -CD complexes with GB, DIMP and DIFP.

Simulant Inclusion in α -CD

Phosphonates with two isopropyl substituents, DIMP and DIFP, whilst they are chemically similar to GB and share its bifurcated hydrogen bonding motif (Figure 4 and Table 6), were unable to penetrate the macrocycle due to the steric bulk of the second

isopropyl group which suggests that they are poor mimics for GB binding in this particular case study. The model is in agreement with a study by Van Hooidek and Breebaart-Hansen in which DIFP was shown to interact with α -CD but 1:1 complex affinity was reported to be weak with K_d increasing from 78 mM at 5 °C to 460 mM at 45 °C (52, 53).

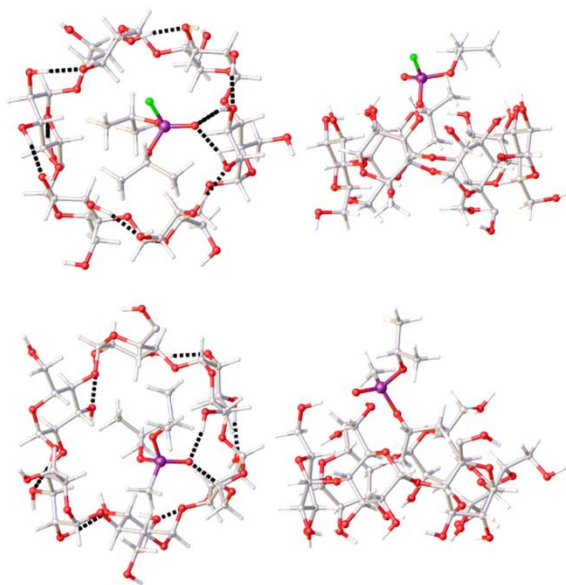


Figure 4. α -CD complexes with (from top) DIFP and DIMP.

DMMP, with a methyl and two methoxy substituents, orients its least bulky substituent towards the cavity in an almost identical manner to GB, and engages in hydrogen bonding between an α -CD hydroxyl group and the phosphonate oxygen (Figure 5 and Table 7). Similar binding is observed in the crystal structure of DMMP with a [4]resorcinarene bearing a single carboxylate group on its upper rim (54). When the [4]resorcinarene carboxylate points into the cavity, DMMP engages in hydrogen bonding and has one methoxy group descending into the macrocycle and shows similarities with the orientation predicted for the hydrogen bond stabilised α -CD complex. DMMP is a widely used simulant for GB and, despite significant structural

differences between the two, the modelling suggests that it is a good mimic and appears to bind to α -CD in a very similar fashion.

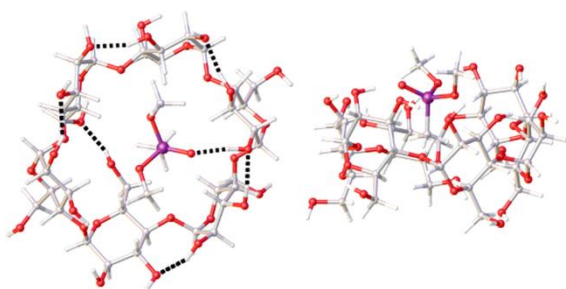


Figure 5. α -CD complexes with DMMP.

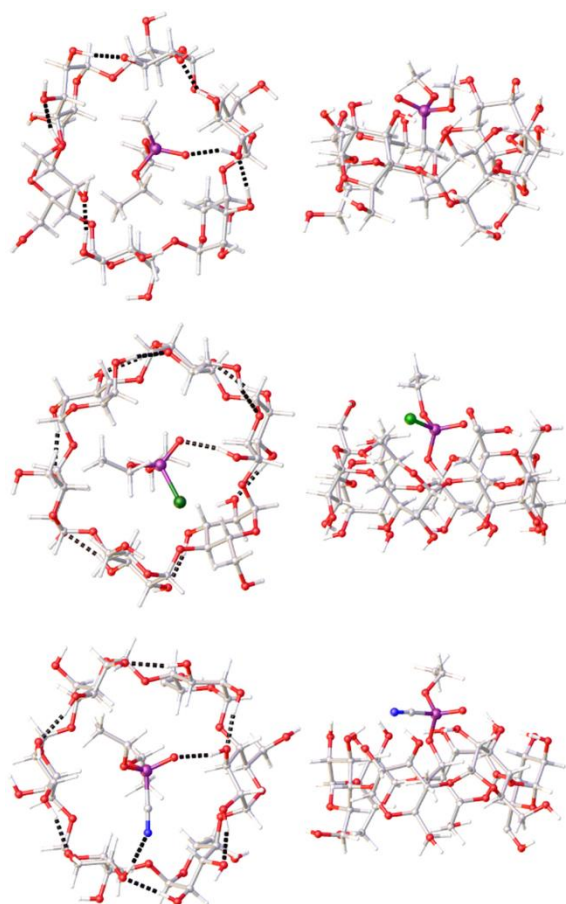


Figure 6. α -CD complexes with (from top) DEMP, DECP and DECyP.

The remaining three complexes with DEMP, DECP and DECyP all incorporate two ethoxy substituents and, as can be seen from Figure 6, one of these is drawn into the α -CD cavity while the phosphonate oxygen engages in hydrogen bonding with the α -

CD hydroxyl group. In all cases the hydrogen bond lengths and angles are similar to the intramolecular interactions found for the cyclodextrin (Table 7), indicating that the simulants are able to compete efficiently with the existing hydrogen bonding network in order to form a stable complex.

Complex	O-H...O		
	O...O(Å)	H...O(Å)	O-H...O(°)
DEMP to α-CD			
O-H...O	2.749	1.836	152.165
α-CD intramolecular			
O-H...O	2.664	1.685	172.373
O-H...O	2.697	1.765	157.005
O-H...O	2.648	1.694	162.079
O-H...O	2.652	1.697	161.908
O-H...O	2.668	1.784	148.143
DMMP to α-CD			
O-H...O	2.764	1.793	168.466
α-CD intramolecular			
O-H...O	2.788	1.859	157.222
O-H...O	2.874	1.904	170.740
O-H...O	2.817	1.901	155.134
O-H...O	2.830	1.861	169.055
O-H...O	2.781	1.949	141.935
O-H...O	2.978	2.092	150.411
DECP to α-CD			
O-H...O	2.853	1.878	164.450
α-CD intramolecular			
O-H...O	2.966	2.017	162.502
O-H...O	2.881	1.933	161.835
O-H...O	2.784	1.955	141.345
O-H...O	2.938	2.003	159.309
O-H...O	2.785	1.856	158.302
O-H...O	2.782	1.822	165.048
O-H...O	2.939	1.998	160.870
O-H...O	2.898	1.977	156.894
DECyP to α-CD		O-H...N	
	O...N(Å)	H...N(Å)	O-H...N(°)
O-H...N	2.835	1.878	165.731
		O-H...O	
	O...O(Å)	H...O(Å)	O-H...O(°)
O-H...O	2.748	1.817	158.502
α-CD intramolecular			
O-H...O	2.918	1.940	174.441
O-H...O	2.772	1.838	158.450
O-H...O	2.783	1.815	167.887
O-H...O	2.995	2.034	167.663
O-H...O	2.804	1.831	169.974
O-H...O	2.830	1.872	165.418

Table 7. Intra- and intermolecular distances (Å) and angles (°) found in simulated α -CD complexes with DEMP, DMMP, DECP and DECyP.

Agents and Simulants

Inspection of the CWAs shows that both the V- and G-series agents contain a chiral phosphorus centre whereas all the simulants are achiral and display pseudo-mirror plane symmetry. The toxicity of the G-series stems from their inactivation of AChE and it has been shown that for the four isomers of GD it is the chirality of the phosphorus, and not that of the carbon, which correlates with toxicity (55, 56). As well as a lack of chirality, the simulants often contain a greater number of bulky substituents than the CWA that they are supposed to mimic. For example, replacing the *P*-methyl group of GB by an isopropoxy group yields diisopropyl fluorophosphate (DIFP). The second isopropoxy group interferes with the compound's ability to reach the active side of AChE resulting in much lower toxicity as borne out by crystal structures of GB and DIFP bound to *Mus musculus* AChE. These show that DIFP can only approach the phosphorylation site through the extensive rearrangement of a 13-residue acyl loop region to accommodate both isopropoxy substituents whereas GB requires no such rearrangement (57). In addition, DIFP has been shown to have inhibition rate constants orders of magnitude lower than the G-series agents GB, GD and GF which reflect the difficulty experienced by DIFP in binding to the active site (58).

A good simulant should bind to a substrate in a similar manner to a CWA while ensuring that other sidechains are oriented away from the binding site. Interestingly, all three simulants with ethoxy substituents, DEMP, DECP and DECyP, are able to thread one ethyl group into the hydrophobic CD cavity and engage in hydrogen bonding. These simulations suggest that the best mimics for GB are not those with isopropoxy substituents but those with ethoxy groups and an electronegative substituent. Thus DEMP, DECP and DECyP appear to be better mimics for GB than DIFP in this specific case study even though the latter has isopropoxy and fluoro substituents in common

with the agent. It is important to note that other factors must also be considered when choosing appropriate simulants for CWAs. For example, DECP hydrolyses readily (59) and, while it may be appropriate as a simulant for gas phase detection systems such as those based on surface plasmon resonance (54) or similar techniques, would be unsuitable if aqueous sampling was undertaken. Similarly, DMMP, despite its attractive lack of steric bulk, typically exhibits quite weak binding (60).

Conclusions

A relatively simple method has been developed to predict the structures of CWAs and simulants with infrared spectra calculated from these structures more closely matching experimental data than previously reported methods. α -CD inclusion complexes of GB and simulants were modelled as the macrocycles are known to bind CWAs (61) and their derivatives have used as agent decontaminants (62). Analysis of these inclusion complexes suggests that simulants with ethoxy substituents, DECP and DECyP, may be able to mimic the inclusion behaviour of GB more effectively than DIMP or DIFP with their two bulky isopropoxy groups. This work complements recent studies into the inclusion behaviour of GD and commonly used simulants within the larger hydrophobic cavity of β -CD where simulant complexation was shown to occur with significantly weaker affinity than GD complexation (63).

Acknowledgments

This work was funded by the US Army Research Office under contract W911NF-15-1-0624. M.R.S. thanks Linda Lee and Oliver Payne (CBR Division, Dstl) for provision of spectral data.

References

1. Convention on the prohibition of the development, production, stockpiling and use of chemical weapons and on their destruction, Technical Secretariat of the Organisation for the Prohibition of Chemical Weapons, 2005 (https://www.opcw.org/fileadmin/OPCW/CWC/CWC_en.pdf accessed September 14th 2017).
2. Black, R. M.; Read, R. W. *Toxin Rev.* **2007**, *26*, 275-298.
3. Morita, H.; Yanagisawa, N.; Nakajima, T.; Shimizu, M.; Hirabayashi, H., Okudera, H.; Nohara, M.; Midorikawa, Y.; Mimura, S. *Lancet* **1995**, *346*, 290-293.
4. Yokoyama, K.; Araki, S.; Murata, K.; Nishikitani, M.; Okumura, T.; Ishimatsu, S.; Takasu, N. *J. Physiology (Paris)* **1998**, *92*, 317-323.
5. Rice, H.; Dalton, C. H.; Graham, S. J.; Green, A. C.; Jenner, J.; Groombridge H. J.; Timperley, C. M. *Proc. Roy. Soc. A* **2015**, *471*, 20140891.
6. Mahato, T. H.; Prasad, G. K.; Singh, B.; Acharya, J.; Srivastava A. R.; Vijayaraghavan, R. *J. Haz. Mat.* **2009**, *165*, 928-932.
7. Lenz, D. E.; Yeung, D.; Smith, J. R.; Sweeney, R. E.; Lumley, L. A.; Cerasoli, D. M. *Toxicology* **2007**, *233*, 31-39.
8. Petrea, N.; Petre, R.; Epure, G.; Şomoghi, V.; Tănase, L. C.; Teodorescu, C. M.; Neaţu, Ş. *Chem. Commun.* **2016**, *52*, 12956-12959.
9. D'Agostino, P. A.; Hancock, J. R.; Chenier, C. L.; Jackson Lepage, C. R. *J. Chromatogr. A* **2006**, *1110*, 86-94.
10. Liang, M.; Fan, K.; Pan, Y.; Jiang, H.; Wang, F.; Yang, D.; Lu, D.; Feng, J.; Zhao, J.; Yang, L.; Yan, X. *Anal. Chem.* **2013**, *85*, 308-312.
11. Dennison, G. H.; Bochet, C. G.; Curty, C.; Ducry, J.; Nielsen, D. J.; Sambrook, M. R.; Zaugg, A.; Johnston, M. R. *Eur. J. Inorg. Chem.* **2016**, 1348-1358.

12. Dennison, G. H.; Sambrook, M. R.; Johnston, M. R. *Chem. Commun.* **2014**, *50*, 195-197.
13. Hiscock, J. R.; Sambrook, M. R.; Cranwell, P. B.; Watts, P.; Vincent, J. C.; Xuereb, D. J.; Wells, N. J.; Raja, R.; Gale, P. A. *Chem. Commun.* **2014**, *50*, 6217-6220.
14. Hiscock, J. R.; Sambrook, M. R.; Wells, N. J.; Gale, P. A. *Chem. Sci.* **2015**, *6*, 5680.
15. Hiscock, J. R.; Sambrook, M. R.; Ede, J. A.; Wells, N. J.; Gale, P. A. *J. Mater. Chem. A* **2015**, *3*, 1230-1234.
16. Ewig, C. S.; Van Wazer, J. R. *J. Mol. Struct.: THEOCHEM* **1985**, *122*, 179-187.
17. Walker, A. R. H.; Suenram, R. D.; Samuels, A.; Jensen, J.; Ellzy, M. W.; Lochner, J. M.; Zeroka, D. *J. Mol. Spectrosc.* **2001**, *207*, 77-82.
18. Suenram, R. D.; Lovas, F. J.; Plusquellic, D. F.; Lesarri, A.; Kawashima, Y.; Jensen, J. O.; Samuels, A. C. *J. Mol. Spec.* **2002**, *211*, 110-118.
19. White, W. E. Technical Report ECBC-TR-265, 2002; Edgewood Chemical Biological Center, RDECOM, U.S. Army.
20. Bermudez, V. M. *J. Phys. Chem. C* **2007**, *111*, 3719-3728.
21. Bermudez, V. M. *J. Phys. Chem. C* **2007**, *111*, 9314-9323.
22. Bermudez, V.M. *Surf. Sci.* **2008**, *602*, 1938-1947.
23. Bermudez, V. M. *J. Phys. Chem. C* **2009**, *113*, 1917-1930.
24. Bermudez, V. M. *J. Phys. Chem. C* **2010**, *114*, 3063-3074.
25. Rivin, D.; Meermeier, G.; Schneider, N. S.; Vishnyakov, A.; Neimark, A. V. *J. Phys. Chem.* **2004**, *B 108*, 8900-8909.
26. Wilmsmeyer, A. R.; Gordon, W. O.; Davis, E. D.; Troya, D.; Mantooth, B. A.; Lalain, T. A.; Morris, J. R. *J. Phys. Chem. C* **2013**, *117*, 15685-15697.

27. Papas, B. N.; Petsalakis, I. D.; Theodorakopoulos, G.; Whitten, J. L. *J. Phys. Chem. C* **2014**, *118*, 23042-23048.
28. Vishnyakov, A.; Neimark, A. V. *J. Phys. Chem. A* **2004**, *108*, 1435-1439.
29. Vishnyakov, A.; Gor, G. Y.; Lee, M. T.; Neimark, A. V. *J. Phys. Chem. A* **2011**, *115*, 5201-5209.
30. Ault, B. S.; Balboa, A.; Tevault, D.; Hurley, M. *J. Phys. Chem. A* **2004**, *108*, 10094-10098.
31. Mott, A. J.; Rez, P. *Spectrochim. Acta A* **2012**, *91*, 256-260.
32. Cuisset, A.; Mouret, G.; Pirali, O.; Roy, P.; Cazier, F.; Nouali, H.; Demaison, J. *J. Phys. Chem. B* **2008**, *112*, 12516-12525.
33. Alam, T. M.; Pearce, C. J. *J. Mol. Model.* **2015**, *21*, 182.
34. Zhang, Y.-P.; Wang, H.-T.; Zhang, W.-P.; Sun, C.; Bai, Y.; Guo, X.-D.; Sun, H. *J. Appl. Spectr.* **2016**, *83*, 673-681.
35. Hameka, H. F.; Carrieri, A. H.; Jensen, J. O. *Phosphorus, Sulfur, Silicon* **1992**, *66*, 1-11.
36. Hameka, H. F.; Carrieri, A. H.; Jensen, J. O. *Int. J. Quantum. Chem.* **1994**, *50*, 161-172.
37. Shan, X.; Vincent, J. C.; Kirkpatrick, S.; Walker, M. D.; Sambrook, M. R.; Clary, D. C. *J. Phys. Chem. A* **2017**, *121*, 6200-6210.
38. *Spartan '16*, Wavefunction Inc., Irvine CA, USA, 2017.
39. Chacko, K. K.; Saenger, W. *J. Am. Chem. Soc.* **1981**, *103*, 1708-1715.
40. Piffath, R. J. Report ECBC-TR-323, **2003**, Edgewood Chemical Biological Center, Aberdeen Proving Ground, MD 21010, USA
41. Sharpe, S. W.; Johnson, T. J.; Chu, P. M.; Kleimeyer, J.; Rowland, B. *Proc. SPIE* **2003**, 5085, 19-27.

42. Heiss, D. R.; Zehnder, D. W., II; Jett, D. A.; Platoff, G. E. Jr.; Yeung D. T.; Brewer, B. N. *J. Chem.* **2016**, 3190891.
43. Sheehan, R.; Cragg, P. J. *Supramol. Chem.* **2008**, 20, 443-451.
44. Miah, M.; Pavey, K. D.; Gun'ko, V. M.; Sheehan, R.; Cragg, P. J. *Supramol. Chem.* **2004**, 16, 185-192.
45. Momany, F. A.; Willett, J. L. *Carbohydr. Res.* **2000**, 326, 210-226.
46. Sha, J.-Q.; Zhong, X.-H.; Wu, L.-H.; Liu, G.-D.; Sheng, N. *RSC Adv.* **2016**, 6, 82977-82983.
47. Liu, Z.; Samanta, A.; Lei, J.; Sun, J.; Wang, Y.; Stoddart, J. F. *J. Am. Chem. Soc.* **2016**, 138, 11643-11653.
48. van Hoodonk, C.; Breebaart-Hansen, J. C. A. E. *Recl. Trav. Pays-Bas* **1970**, 89, 289-299.
49. Sirin, G. S.; Zhou, Y.; Lior-Hoffmann, L.; Wang, S.; Zhang, Y. *J. Phys. Chem. B* **2012**, 116, 12199-12207.
50. Sirin, G. S.; Zhang, Y. *J. Phys. Chem. A* **2014**, 118, 9132-9139.
51. Désiré, B.; Saint-André, S. *Experientia* **1987**, 43, 395-397.
52. van Hoodonk, C.; Breebaart-Hansen, J.C.A.E. *Recl. Trav. Pays-Bas* **1971**, 90, 680-686.
53. Poh, B.-L.; Saenger, W. *Spectrochim. Acta* **1983**, 39A, 305-307.
54. Daly, S. M.; Grassi, M.; Shenoy, D. K.; Ugozzoli, F.; Dalcanale, E. *J. Mater. Chem.* **2007**, 17, 1809-1818.
55. Benschop, H. P.; De Jong, L. P. A. *Acc. Chem. Res.* **1988**, 21, 368-374.
56. van Dongen, C. J.; de Lange, J.; van Genderen, J. *Biochem. Pharmacol.* **1989**, 38, 2263-2268.
57. Hörnberg, A.; Tunemalm, A.-K.; Ekström, F. *Biochemistry* **2007**, 46, 4815-4825.

58. Worek, F.; Thiermann, H.; Szinicz, L.; Eyer, P. *Biochem. Pharmacol.* **2004**, *68*, 2237-2248.
59. Alam, T. M.; Kinnan, M. K.; Wilson, B. W.; Wheeler, D. R. *ChemistrySelect* **2016**, *1*, 2698-2705.
60. Hiscock, J. R.; Wells, N. J.; Ede, J. A.; Gale, P. A.; Sambrook, M. R. *Org. Biomol. Chem.*, **2016**, *14*, 9560.
61. Letort, S.; Balieu, S.; Erb, W.; Gouhier, G.; Estour, F. *Beilstein J. Org. Chem.* **2016**, *12*, 204-228.
62. Brandhuber, F.; Zengerle, M.; Porwol, L.; Bierwisch, A.; Koller, M.; Reiter, G.; Worek, F.; Kubik, S. *Chem. Commun.* **2013**, *49*, 3425-3427.
63. Sambrook, M. R.; Vincent, J. C.; Ede, J. A.; Gass, I. A.; Cragg, P. J. *RSC Adv.* **2017**, *7*, 38069.