This is a pre-print manuscript of an article published in the journal *Pure and Applied Chemistry*.

The published version of this article is available at the De Gruyter website:
https://www.degruyter.com/document/doi/10.1515/pac-2022-1113/html

How to cite: Costanzi, Stefano, Koblentz, Gregory D. and Cupitt, Richard T. "Expanding the Australia Group's chemical weapons precursors control list with a family-based approach" Pure and Applied Chemistry, 2023. https://doi.org/10.1515/pac-2022-1113

Pure Appl. Chem.

Expanding the Australia Group's Chemical Weapons Precursors Control List with a Family-Based Approach

Stefano Costanzi,[a‡] [a] Gregory D. Koblentz,[b] and Richard T. Cupitt[c]

[a]*Department of Chemistry, American University, 4400 Massachusetts Avenue, NW, Washington, DC 20016, USA* [b]*Schar School of Policy and Government, George Mason University, 3351 Fairfax Drive, Arlington, VA 22201, USA* [c]*The Henry L. Stimson Center, 1211 Connecticut Ave, NW, Washington, DC 20036, USA*

Abstract: The Australia Group (AG) is a forum of like-minded states seeking to harmonize export controls to prevent the proliferation of chemical and biological weapons. The AG Chemical Weapons Precursors list features dual-use chemicals that can be used as precursors for the synthesis of chemical weapons, all individually enumerated. This is in contrast with the Chemical Weapons Convention (CWC) Schedules, which, alongside entries describing discrete chemicals, also include entries that describe families of chemicals. By using families of chemicals, the CWC achieves the objective of covering with a single entry a wide array of related chemicals of concern, including chemicals that have not yet been made. There are practical reasons why the AG Chemical Weapons Precursors list is exclusively based on the enumeration of individual chemicals. A cheminformatics tool of which we have developed a prototype, the Nonproliferation Compliance Cheminformatics Tool (NCCT), has the potential to enable export control officers to handle control lists that contain families of chemicals. Thus, it opens the way to expand the AG Chemical Weapons Precursors list to a family-based approach for some of its entries. Such a change would result in a closer alignment of the chemical space covered by the AG Chemical Weapons Precursors list with that covered by the CWC Schedules, thus closing loopholes that could be exploited by proliferators.

Keywords: cheminformatics, chemical weapons, precursors, export controls, Australia Group (AG), Chemical Weapons Convention (CWC).

INTRODUCTION

‡ Corresponding author. Email: costanzi@american.edu

Chemical weapons are weapons that exploit the toxicity of chemicals to kill or harm humans or animals [1]. A complete ban on chemical weapons is posed by the Chemical Weapons Convention (CWC), an international treaty that entered into force in 1997 that enjoys almost universal embracement. This comprehensive ban prohibits not only the use of chemical weapons, but also their development, production, stockpiling, retaining, and transfer [2]. Members are also prohibited from providing assistance to chemical weapons programs in other countries. However, despite the CWC, chemical weapons are far from being a relic of the past. In recent years, they have been used by state and non-state actors in the middle east, particularly in the context of the Syrian Civil War. They have also been used in multiple occasions for targeted assassinations and assassination attempts [3,4].

Established in 1985, the Australia Group (AG) is an informal multilateral forum of 42 participant countries, plus the European Union (EU), intended to harmonize export control legislation to prevent the proliferation of chemical and biological weapons [5,6]. The AG has developed its Common Control Lists (CCL), which all participants agree to apply to their export licensing measures by incorporating them into their national export control lists. Among the CCL is the AG Chemical Weapons Precursors list (**Fig. 1**) [7–9]. The AG is a concrete manifestation of the CWC's ban on aiding chemical weapons programs in other states.

In the following paragraphs, we illustrate how the AG Chemical Weapons Precursors list can be made more comprehensive by adopting a family-based approach for the listing of some of the chemicals of concern. We also explain how cheminformatics solutions can enable such a change, thus fostering a closer alignment between the AG Chemical Weapons Precursors list and the CWC Schedules and ultimately helping close gaps that could be exploited by proliferators.

FAMILIES OF CHEMICALS VS INDIVIDUAL CHEMICALS

In its current form, the AG Chemical Weapons Precursors list comprises 89 individually enumerated chemicals. Each entry on the list describes an individual chemical, identified through its chemical name as well as the registry number assigned to it by the Chemical Abstract Service (CAS) division of the American Chemical Society (CAS Registry Number®, or CAS RN®) [7,9].

This approach is different from the one taken by the three schedules of chemicals contained in the Annex on Chemicals of the CWC (CWC Schedules). These are lists of toxic chemicals and precursors for their synthesis whose purpose is to support the CWC verification regime and declaration requirements. Unlike the AG Chemical Weapons Precursors list, the CWC Schedules include entries that describe families of chemicals, defined by a scaffold with variable chemical groups attached to it, as well as entries describing discrete chemicals that sometimes appear as examples of chemicals within a larger chemical group. By using families of chemicals, the CWC achieves the objective of covering with a single entry a wide array of related chemicals of concern, including chemicals that have not yet been made [8–10]. For instance, the first entry of CWC Schedule 1 covers over two million chemicals structurally related to the nerve agents sarin and soman, which are listed as specific examples (**Fig. 2**) [11,12]. As explained more thoroughly in the "**Families can be problematic**" section, the AG and other lists rely on individually enumerated chemicals for the relative ease for licensing and enforcement of export controls by non-chemists, who would not be able to assess whether a chemical under scrutiny falls within the scope of a family of chemicals.

Below, we provide two case studies that, by comparing related entries in the AG Chemical Weapons Precursors list and the CWC Schedules, illustrate the differences between the coverage provided by entries

that describe individual chemicals versus that provided by entries that describe families of chemicals. These case studies illustrate the value of the more comprehensive, family-based approach embodied in the CWC Schedules for reducing loopholes that proliferators can exploit to acquire chemical weapons. In particular, shifting to a family-based approach for some of the entries of the AG Chemical Weapons Precursors would allow for a closer alignment of the chemical space that it covers with the one covered by the CWC Schedules, and more accurately address the associated risk.

The first case study involves chloroethylamines that can be employed as precursors for the synthesis of nerve agents of the V series, including VX, VM, and VR. Entry 2B10 of CWC Schedule 2 covers a family of chloroethylamines that, among other members, comprises the one that can be used as a precursor for the synthesis of VX. The possible combinations of the two alkyl substituents account for 10 different molecules. However, considering that each of these 10 molecules can exist in multiple variants (e.g. salts with different counterions, isotopically labeled variants, etc.), there are 84 known chemicals that belong to this family, as well as other chemicals that have never been reported but that could conceivably be made in the future (**Fig. 3A**) [11,12]. Conversely, the AG Chemical Weapons Precursors list, as entries 11 and 54, covers exclusively a specific chloroethylamine that can be used as a precursor for the synthesis of the nerve agent VX, in its free form as well as its hydrochloride salt (**Fig. 3B**). As this case demonstrates, the approach taken by the CWC Schedules is more comprehensive than the one taken by the AG. For instance, a precursor for the synthesis of nerve agents VM and VR is covered by the Entry 2B10 of CWC Schedule 2, but not by the AG Chemical Weapons Precursors list (**Fig. 3C**) [13]. When Syria acceded to the CWC in 2013, it declared the possession of 25 tons of the chloroethylamine that can be used for the synthesis of VM and VR, possibly acquired or produced as a consequence of the fact that it had encountered difficulties in procuring precursors for the preparation of VX. That Syria was able to identify a synthesis pathway that avoided the use of an AG-controlled chemical to produce a V-series nerve agent is concerning. As this example illustrates, adopting a family-based approach to chemical control lists allows for a much wider coverage than listing individual chemicals, thus covering dangerous loopholes that would remain open by listing exclusively individual chemicals [13].

The second case study involves amidines that can be employed as precursors for nerve agents of the Novichok series [14,15]. In March of 2018, an assassination attempt against Sergei Skripal was conducted in Salisbury, UK. The amidine-bearing Novichok agent used in the Salisbury incident was not covered by the CWC Schedules. Similarly, direct precursors for their synthesis were not covered by either the CWC Schedules or the AG Chemical Weapons Precursors list [16,17]. An amendment of the CWC Schedules approved by consensus by the Organisation for the Prohibition of Chemical Weapons (OPCW)'s Conference of State Parties in December 2019 led to the addition to CWC Schedule 1 of four new entries, two of which cover two large families of amidine-bearing Novichoks (CWC 1A13 and CWC 1A14) (**Fig 4A**). Precursors for the synthesis of these chemicals were not added to the CWC Schedules. However, soon after the CWC Schedules amendment, 24 individually listed chemical precursors for the synthesis of Novichok agents, including 17 amidines, were added to the AG Chemical Weapons Precursors list (**Fig 4B**). This revision, however, did not fully capture the range of amidine precursors that could be used to produce the amidine-bearing Novichok agents covered by these two entries. Notably, in entries CWC 1A13 and CWC 1A14 of CWC Schedule 1, the amidine branches of the Novichok agents are defined very broadly as alkyl or cycloalkyl groups each with a size of up to 10 carbon atoms. Conversely, the amidines individually listed in the AG Chemical Weapons Precursors list feature only alkyl groups with up to three carbon atoms. It is also worth noting that, even within this reduced subset, not all combinations of alkyl chains with up to three carbon atoms have been listed. Four of the missing combinations identify chemicals

that have never been reported in the literature and, consequently, do not have a CAS Registry Numbers associated with them. However, the fact that these chemicals have never been reported does not mean that they cannot be produced and exported in the future. As we recently proposed, using a family-based approach to define these Novichok precursors in the AG Chemical Weapons Precursors list, in a way that mirrors the CWC Schedules, would provide a wider coverage to these chemicals and close gaps that could be exploited by proliferators [18]. As indicated in the latest version of the AG CCL Handbook, these amidines do not have significant commercial applications and are typically used in small quantities for research purposes [19]. Hence, they could be used for a pilot study to test the feasibility of extending the family-based approach to the AG Chemical Weapons Precursors list. This case study also illustrates another important point: given that the AG is a forum of a relatively small number of like-minded countries, the AG Chemical Weapons Precursors list can be amended more easily than the CWC Schedules. Hence, enabling the inclusion of families of chemicals in the AG Chemical Weapons Precursors list would allow a rapid coverage of a large number of chemicals by export control lists before (or in the absence of) an amendment of the CWC Schedules.

FAMILIES CAN BE PROBLEMATIC

Despite the significant benefits in terms of coverage provided by a family-based approach, working with families of chemicals is not easy and there are practical reasons why the AG Chemical Weapons Precursors list is exclusively based on the enumeration of individual chemicals. For instance, sticking with the example of choroethylamines, given that there are no listed chemical names or registry numbers in entry 2B10 of CWC Schedule 2 (**Fig. 3**), it would be virtually impossible for a frontline officer without sufficient chemical knowledge to gauge whether the precursor for the synthesis of VM falls within the scope of this entry based on its name (2-Chloro-N,N-diethylethanamine) or its CAS Registry Number (100-35-6) [13].

The different scope and purposes of the AG and the CWC must be kept in mind [8]. The AG Chemical Weapons Precursors list, although developed by individuals with a significant level of chemistry training, is intended to be used by non-chemists as well. Beyond chemical manufactures and exporters, the list is intended to be used by export control officers, who typically lack a significant background in chemistry and would be unable to gauge whether a chemical under scrutiny falls within the scope of a family of chemicals. They also work under significant time constraints with a focus on trade facilitation. Conversely, the CWC Schedules, which are intended to support the treaty's verification regime, are geared towards inspectors and officials from the OPCW, who have substantial chemistry knowledge and are capable of interpreting families of chemicals.

As outlined below, the adoption of cheminformatics tools could open the way for the inclusion of families of chemicals in the AG Chemical Weapons Precursors list, by enabling frontline officers to handle this sort of entry. Cheminformatics can provide effective technical solutions to the practical issues that currently prevent a wider use of families of chemicals in lists of chemicals of security concern. Political commitment will also be needed to fully leverage these solutions.

BRINGING FAMILIES TO THE AG WITH CHEMINFORMATICS

Cheminformatics, a discipline at the intersection of chemistry and information technology, provides an opportunity to enable individuals with limited chemistry knowledge to work with families of chemicals. As

a result, it opens the door to a wider way of employing and managing families of chemicals, even in lists that, like the AG Chemical Weapons Precursors list, have not traditionally taken this approach. To illustrate and test this concept, we have developed a working prototype of a cheminformatics database that automates the task of assessing whether a chemical under scrutiny is covered by a given control list. We call this database the Nonproliferation Cheminformatics Compliance Tool (NCCT). As we have previously described, we have developed the NCCT database prototype based on Instant JChem (IJC), a commercial software for the implementation, management, and query of databases of chemical structures produced by the company ChemAxon [20,21].

The NCCT database automates the process of checking whether a chemical under scrutiny is covered by one or more of the lists embedded in the database, thus overcoming the major problems associated with the manual consultation of the lists, and it does so within seconds, which addresses the considerable time pressures facing export and border control officers. Although this article focuses on the problems posed by families of chemicals, it should be noted that entries that describe individual chemicals pose challenges too, especially due to the myriad of synonyms with which the same chemical can be called and the fact that different variants of the same chemical are assigned different CAS Registry Numbers. The NCCT database helps addressing those issues as well [20].

We have currently incorporated into the NCCT database lists from a variety of frameworks for chemical disarmament and nonproliferation. In addition to the CWC Schedules and the AG Chemical Weapons Precursors list, the NCCT includes the Wassenaar Arrangement (WA) Munitions List 7 (ML7), the EU's Syria-related list, and the strategic chemicals identified in the World Customs Organization (WCO) Strategic Trade Control Enforcement Implementation Guide (STCE). Some of these lists, like the CWC Schedules and the WA ML7, feature both individual chemicals and families of chemicals. Conversely, other lists, like the AG Chemical Weapons Precursors list, EU Syria-related list, and the WCO STCE, feature exclusively individual chemicals [20]. These lists have been implemented for development and testing purposes. For their deployment in a specific country, the cheminformatics databases would have to be tailored to the demands of different legal requirements of national jurisdictions related to compliance with and enforcement of export control legislation and regulations.

The NCCT shifts the focus from names and CAS Registry Numbers to chemical structures, which are better suited to describing chemicals. Through the IJC interface, an operator can input a query chemical in a variety of different ways, including by entering its chemical name, CAS Registry Number, or a structural identifier, which is a string that encodes the structure of the chemical. Alternatively, an operator can choose to draw the structure of the chemical, which is a feature mostly relevant to those who have sufficient background in chemistry [20].

IJC converts the input—whether a name, CAS Registry Number, or structural identifier—into a chemical structure and subsequently standardizes it to establish the equivalence between all variants of the same chemical. The standardized chemical is then checked against our standardized database of chemical weapons control lists. The database entries that match the query are then pulled out. The key feature of the NCCT database is that it contains not only the structures of the individually listed chemicals, but also the generic structures that describe the listed families of chemicals. It is this feature that allows the matching of a query chemical with a family of chemicals that covers it [20].

To practically illustrate the functioning of the NCCT, it is worth reexamining the case study of the chloroethylamines that are precursors for the synthesis of V-series nerve agents (**Fig. 3**). As mentioned, given that there are no listed chemical names or CAS Registry Numbers in entry 2B10 of CWC Schedule 2, it would be virtually impossible for a frontline officer without sufficient training in chemistry to gauge

whether a given chemical, such as the chloroethylamine that can be used as a precursor for VX, falls within the scope of this entry based on its name or CAS Registry Number. The NCCT solves this issue by automating the process (**Fig. 5**). A frontline officer enters the chemical name or the CAS Registry Number into the IJC interface and IJC converts it into a chemical structure. The structure is then used to run a search against the database and, after a few seconds, the tool displays for each of the lists of chemicals implemented in the database the number of entries that match the query chemical (hits). A click on the hits links pulls up the entries that match the query (**Fig. 6**). Notably, the chemical name or the CAS Registry number entered into the NCCT do not need to be the same as those listed in the CW control lists, as the NCCT is a database of chemical structures, not of chemical names and registry numbers. As long as the ICJ engine is capable of converting the entered chemical name or CAS Registry Number into a structure, then the entered chemical can be successfully checked against the NCCT database. However, as we explain in the next section, there are limits to the conversion capabilities of IJC.

In the case of the VX chloroethylamine precursor, the results indicate that the query chemical is covered by three of the lists included in the NCCT database. Namely, it is covered by CWC Schedule 2, as a member of the family of chemicals listed under entry 2B10, as well as an individual chemical by the AG Chemical Weapons Precursors list and the WCO STCE. For individual chemicals the entries are annotated with additional information, including for instance a link to the relevant PubChem entry, where one can find a wealth of information on the chemical in question, including toxicity and safety data.

TOWARDS THE DEVELOPMENT OF DEPLOYABLE CHEMINFORMATICS TOOLS

The prototype cheminformatics database that we have developed and that is currently undergoing field testing is intended as a model tool to illustrate how cheminformatics can support the work of individuals who deal with lists of controlled chemicals, in particular those who work in the fields of export controls and customs. Through field tests and feedback from stakeholders, we have probed the limits of the prototype, some of which are listed below. Based on the lessons learned while testing the NCCT prototype, the next goal is to foster the development of cheminformatics tools that overcome these limits and are suitable for deployment in the field. Some of the requirements for the development of these tools are listed below.

First, the IJC software used for the implementation and query of the NCCT database is rather complex and has an interface designed with chemists in mind [20]. Going forward, deployable cheminformatics tools will need to have a simpler, web-based interface; they will need to have the ability to enter multiple queries in batch and, possibly, an optical recognition system that allows the scanning of chemical information directly into the system; moreover, they will need the ability to recognize and extract relevant chemical information from strings of text that contain additional information (such as information on quantity, concentration, packaging etc.). Controlled Substances Squared, by Scitegrity, and Compliance Checker, by ChemAxon, are very good models and potential platforms for such developments. As we explained elsewhere, both of these tools, whose primary focus is controlled substances of abuse, can handle families of chemicals and are endowed with a relatively simple, web-based interface [22–24]. The Chemical Abstract Service (CAS) databases, which already feature regulatory information for the listed chemicals, could be another potential development platform. The tool will have to be designed with cybersecurity measures to prevent the hacking or spoofing of the database libraries and to ensure the confidentiality of the searches.

Another important limit is that the NCCT relies entirely on ChemAxon's IJC engine to convert names and CAS Registry Numbers to molecular structures. However, some names and some CAS Registry Numbers cannot be converted into chemical structures by IJC [20]. Going forward, deployable cheminformatics tools will need to be endowed with a very robust engine for the conversion of names and CAS Registry Numbers to molecular structures. Moreover, a thorough annotation of the database with searchable metadata that includes synonyms and CAS Registry Numbers of known variants is desirable.

STRUCTURAL IDENTIFIERS AND SAFETY DATA SHEETS

Beyond names and CAS Registry Numbers, structural identifiers, such as a SMILES or an InChI code, are well-suited for entering chemicals into cheminformatics tools [25–27]. Hence, in conjunction with the development of cheminformatics tools, increasing the availability of structural identifiers to export control officers and customs officials is highly desirable. Structural identifiers are text strings that define molecular structures. They can be algorithmically converted into other structural formats by cheminformatics software [9]. For instance, while neither the name nor the CAS Registry Number for an isotopically labeled version of the VX chloroethylamine precursor can be converted into a structure by IJC, its structural identifiers can (**Fig. 7**). Hence, unlike its name or CAS Registry Number, the structural identifiers of the VX precursor can be used to query the NCCT, leading to the same results that would be obtained when searching for the unlabeled version of the chemical. This is very important given the recommendation of the Scientific Advisory Board (SAB) to the OPCW according to which "the isotopically labelled compound or stereoisomer related to the parent chemical specified in the schedule should be interpreted as belonging to the same schedule" [28–30]. A key issue is that customs officials and export control officers do not typically have access to structural identifiers for the chemicals under scrutiny. However, local and regional regulations widely require chemical manufacturers, distributors, and importers to provide Safety Data Sheets (SDS) for hazardous chemicals [31,32]. The information that goes into the SDS is set by the UN Global Harmonized System for the Classification and Labelling of Chemicals (GHS), according to which chemicals should be identified in SDS through chemical names, CAS Registry Numbers, or other unique identifiers [33]. Adding structural identifiers to the GHS requirements for the identification of discrete chemicals of known chemical structure in SDS would provide frontline officers with access to structural identifiers for all hazardous chemicals. For instance, whenever the identity of a chemical is unclear, export control officers or customs officials could retrieve the structural identifier from the accompanying SDS (which they could find inside the package or could request from the exporter). Similarly, it would be useful to include such information in the prospective European Union chemical digital passports. In turn, this would make it easier to leverage the power of cheminformatics to automate the process of checking lists of controlled chemicals.

CONCLUSIONS

Unlike the CWC Schedules, which feature both individually enumerated chemicals as well as families of chemicals, the AG Chemical Weapons Precursors list is exclusively based on the individual enumeration of chemicals. As a result, the chemical weapons precursors covered by the AG account only for a limited subset of the corresponding agents and precursors listed in the CWC Schedules. This approach leaves gaps open that could be exploited by proliferators [9,13]. Although this article focuses on the AG Chemical

Weapons Precursors list, the same considerations apply to other international or country specific lists that are solely based on the enumeration of individual chemicals.

There are practical reasons why the AG Chemical Weapons Precursors list is exclusively based on the enumeration of individual chemicals [8]. Above all, the AG Chemical Weapons Precursors list is geared towards export control officers, who typically do not have extensive training in chemistry and would be unable to gauge whether a chemical under scrutiny falls within the scope of a family of controlled chemicals. Conversely, the CWC Schedules, which are intended to support the treaty's verification regime, are geared towards inspectors and officials from the OPCW, who have substantial chemistry knowledge and are capable of interpreting families of chemicals.

A cheminformatics database of which we have developed a functioning prototype, the NCCT, has the potential of enabling non-chemists export control officers to handle control lists that contain families of chemicals. The NCCT automates the task of checking whether a chemical under scrutiny is covered by a control list, either because it is explicitly listed as a discrete chemical or because it falls within the scope of a listed family of chemicals [20]. Hence, the NCCT opens the way to expand the AG Chemical Weapons Precursors list by adopting, for some of its entries, a family-based approach, thus closing loopholes that could be exploited by proliferators. For instance, while the CWC Schedules cover a wide range of amidine-bearing Novichok agents, the AG Chemical Weapons Precursors list covers the precursors for the synthesis of only a small subset of them. This gap could be closed by adding to the AG Chemical Weapons Precursors list families of Novichok precursors that mirror those used in the CWC Schedules to describe the agents [18].

Due to the implications that adding chemicals to the AG Chemical Weapons Precursors list may have for the adherent countries and the chemical industry, any revision should be crafted to minimize its impact on the peaceful applications of chemistry, while at the same time maximizing its benefits for chemical weapons disarmament and nonproliferation. This is especially important for precursor chemicals, which are often dual-use items that, beyond their role in the production of chemical warfare agents, may also have legitimate applications [16,18]. Hence, when considering the addition of families of chemicals to control lists, it is important to make sure that these are large enough to include all chemicals of concern, but not so large that their coverage extends to chemicals of low security risk.

RESEARCH FUNDING

This work was supported by Global Affairs Canada under award CWC-2022-0002 (P0011281).

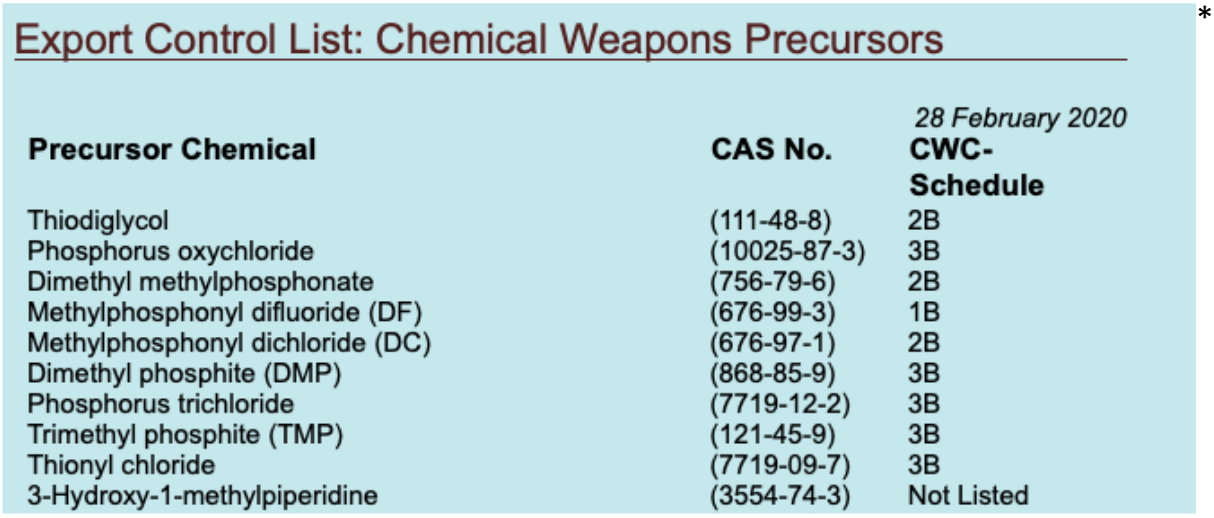
REFERENCES

[1] S. Costanzi, in *Kirk-Othmer Encyclopedia of Chemical Technology* (Wiley Online Library, 2020), pp. 1–32.
[2] OPCW, The Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction, https://www.opcw.org/chemical-weapons-convention.
[3] R. K. Hersman and W. Pittinos, *Restoring Restraint: Enforcing Accountability for Users of Chemical Weapons* (Rowman & Littlefield, 2018).
[4] R. K. Hersman and S. Claeys, *Rigid Structures, Evolving Threat: Preventing the Proliferation and Use of Chemical Weapons* (Center for Strategic & International Studies, 2019).
[5] The Australia Group, The Australia Group: An Introduction, https://www.dfat.gov.au/publications/minisite/theaustraliagroupnet/site/en/precursors.html.
[6] J. I. Seevaratnam, The Nonproliferation Review **13**, 401 (2006).

[7] The Australia Group, Chemical Weapons Precursors, https://www.dfat.gov.au/publications/minisite/theaustraliagroupnet/site/en/precursors.html.
[8] R. J. Mathews, Chemical Weapons Convention Bulletin 1 (1993).
[9] S. Costanzi, C. K. Slavick, B. O. Hutcheson, G. D. Koblentz, and R. T. Cupitt, J. Chem. Inf. Model. **60**, 4804 (2020).
[10] OPCW, Annex on Chemicals, https://www.opcw.org/chemical-weapons-convention/annexes/annex-chemicals/annex-chemicals.
[11] G. Pontes, J. Schneider, P. Brud, L. Benderitter, B. Fourie, C. Tang, C. M. Timperley, and J. E. Forman, Journal of Chemical Education **97**, 1715 (2020).
[12] C. M. Timperley and J. E. Forman, Journal of Chemical Education **98**, 1468 (2021).
[13] S. Costanzi, G. D. Koblentz, and R. T. Cupitt, Strategic Trade Review **6**, 69 (2020).
[14] M. Kloske and Z. Witkiewicz, Chemosphere **221**, 672 (2019).
[15] S. Costanzi, J.-H. Machado, and M. Mitchell, ACS Chem. Neurosci. **9**, 873 (2018).
[16] S. Costanzi and G. D. Koblentz, The Nonproliferation Review **26**, 1 (2019).
[17] S. Costanzi and G. D. Koblentz, Arms Control Today **50**, 16 (2020).
[18] S. Costanzi and G. D. Koblentz, The Nonproliferation Review **0**, 1 (2022).
[19] The Australia Group, Australia Group Common Control List Handbook, https://www.dfat.gov.au/sites/default/files/australia-group-common-control-list-handbook-volume-i.pdf.
[20] S. Costanzi, C. K. Slavick, J. M. Abides, G. D. Koblentz, M. Vecellio, and R. T. Cupitt, Pure and Applied Chemistry **94**, 783 (2022).
[21] ChemAxon, Instant IChem, https://chemaxon.com/products/instant-jchem.
[22] D. Taylor, S. G. Bowden, R. Knorr, D. R. Wilson, J. Proudfoot, and A. E. Dunlop, Drug Discovery Today **20**, 175 (2015).
[23] Scitegrity, Controlled Substances Squared, https://scitegrity.co.uk/index.php?page=cs2.
[24] ChemAxon, Compliance Checker, https://chemaxon.com/products/compliance-checker.
[25] D. Weininger, J. Chem. Inf. Comput. Sci. **28**, 31 (1988).
[26] D. Weininger, A. Weininger, and J. L. Weininger, Journal of Chemical Information and Computer Sciences **29**, 97 (1989).
[27] S. R. Heller, A. McNaught, I. Pletnev, S. Stein, and D. Tchekhovskoi, Journal of Cheminformatics **7**, 23 (2015).
[28] C. M. Timperley, J. E. Forman, M. Abdollahi, A. S. Al-Amri, I. P. Alonso, A. Baulig, V. Borrett, F. A. Cariño, C. Curty, D. Gonzalez, Z. Kovarik, R. Martínez-Álvarez, R. Mikulak, N. M. Fusaro Mourão, P. Ramasami, S. Neffe, S. K. Raza, V. Rubaylo, K. Takeuchi, C. Tang, F. Trifirò, F. M. van Straten, P. S. Vanninen, V. Zaitsev, F. Waqar, M. S. Zina, S. Holen, and H. A. Weinstein, Pure and Applied Chemistry **90**, 1647 (2018).
[29] OPCW, Response to the Director-General's request to the Scientific Advisory Board to provide further advice on scheduled chemicals, SAB-23/WP.1 (2016), https://www.opcw.org/sites/default/files/documents/CSP/RC-4/en/rc4dg01_e_.pdf.
[30] OPCW, Report of The Scientific Advisory Board on Developments in Science and Technology for the Fourth Special Session of the Conference of the States Parties to Review Figure the Operation of the Chemical Weapons Convention, RC-4/DG.1, 30-31 (2018), https://www.opcw.org/sites/default/files/documents/CSP/RC-4/en/rc4dg01_e_.pdf.
[31] International Trade Administration, Special Documents used in Exporting, https://www.trade.gov/special-documents.
[32] OSHA, Hazard Communication Standard: Safety Data Sheets, https://www.osha.gov/sites/default/files/publications/OSHA3514.pdf.
[33] UNECE. Globally Harmonized System of Classification and Labelling of Chemicals (GHS) - Eighth revised edition, p. 37, https://unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev08/ST-SG-AC10-30-Rev8e.pdf.

[34] OPCW, Annex on Chemicals - Schedule 1, https://www.opcw.org/chemical-weapons-convention/annexes/annex-chemicals/schedule-1.

FIGURES AND CAPTIONS



Export Control List: Chemical Weapons Precursors

28 February 2020

Precursor Chemical	CAS No.	CWC-Schedule
Thiodiglycol	(111-48-8)	2B
Phosphorus oxychloride	(10025-87-3)	3B
Dimethyl methylphosphonate	(756-79-6)	2B
Methylphosphonyl difluoride (DF)	(676-99-3)	1B
Methylphosphonyl dichloride (DC)	(676-97-1)	2B
Dimethyl phosphite (DMP)	(868-85-9)	3B
Phosphorus trichloride	(7719-12-2)	3B
Trimethyl phosphite (TMP)	(121-45-9)	3B
Thionyl chloride	(7719-09-7)	3B
3-Hydroxy-1-methylpiperidine	(3554-74-3)	Not Listed

*** Snippet of the AG Chemical Weapons Precursors list from the AG website.**

Fig. 1. A snippet of the AG Chemical Weapons Precursors list from the AG website, showing the first ten entries of the list [7]. The entries identify precursors for the synthesis of nerve agents, vesicants, and psychoactive chemicals. A structurally annotated version of the list is available at the Costanzi Research Website (https://costanziresearch.com/cw-nonproliferation/cw-control-lists/ag-chemicals/) [9].

*

		(CAS registry number)
(1)	O-Alkyl (<=C10, incl. cycloalkyl) alkyl (Me, Et, n-Pr or i-Pr)-phosphonofluoridates	
e.g.	Sarin: O-Isopropyl methylphosphonofluoridate	(107-44-8)
	Soman: O-Pinacolyl methylphosphonofluoridate	(96-64-0)
(2)	O-Alkyl (<=C10, incl. cycloalkyl) N,N-dialkyl (Me, Et, n-Pr or i-Pr) phosphoramidocyanidates	
e.g.	Tabun:O-Ethyl N,N-dimethyl phosphoramidocyanidate	(77-81-6)
(3)	O-Alkyl (H or <=C10, incl. cycloalkyl) S-2-dialkyl (Me, Et, n-Pr or i-Pr)-aminoethyl alkyl (Me, Et, n-Pr or i-Pr) phosphonothiolates and corresponding alkylated or protonated salts	
e.g.	VX: O-Ethyl S-2-diisopropylaminoethyl methyl phosphonothiolate	(50782-69-9)

*** Snippet of CWC Schedule 1 from OPCW website.**

Fig. 2. A snippet of CWC Schedule 1 from the OPCW website, showing the first three entries of the list [34]. The entries identify three large families of nerve agents and provide, sarin, soman, tabun, and VX as specific examples. A structurally annotated version of the list is available at the Costanzi Research Website (https://costanziresearch.com/cw-nonproliferation/cw-control-lists/cwc-schedules/) [9].

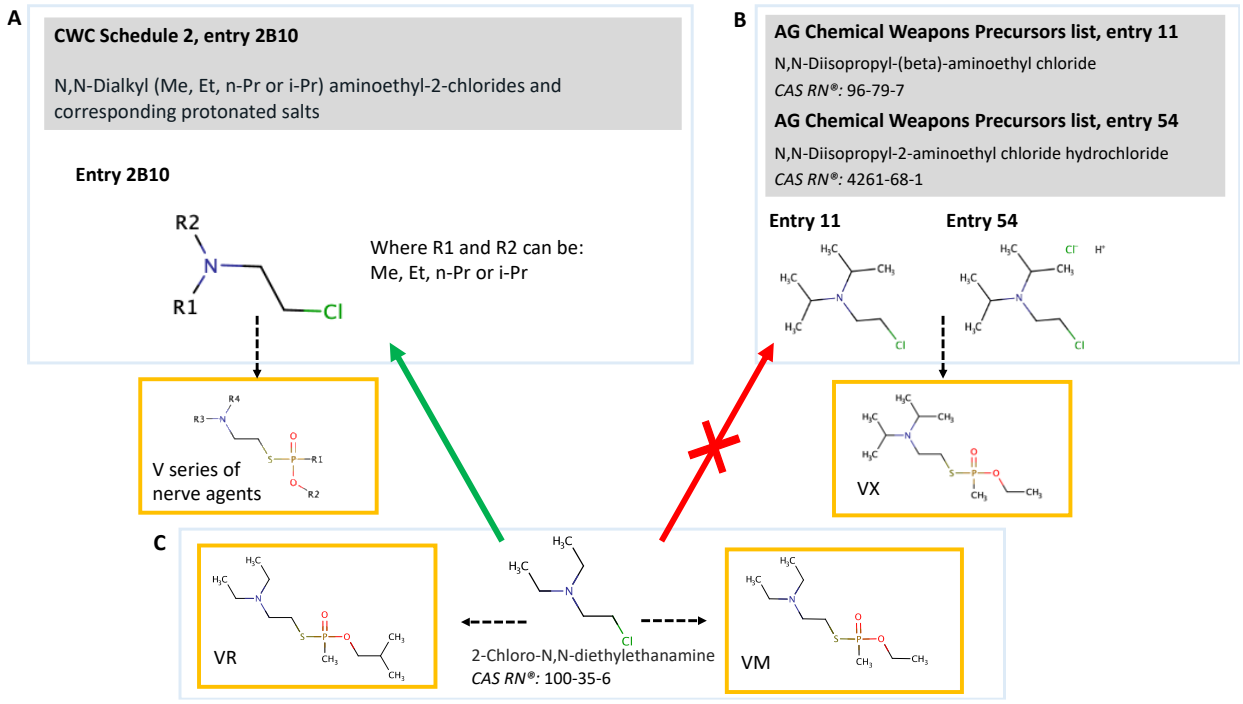
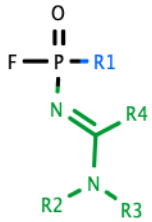


Fig. 3. Panel A: CWC Schedule 2 covers a family of chloroethylamines that can be used for the synthesis nerve agents of the V-series. Panel B: the AG Chemical Weapons Precursors list covers exclusively a single chloroethylamine precursor, namely the one that can be used for the synthesis of the nerve agent VX. Panel C: the chloroethylamine precursor for the synthesis of the nerve agents VM and VR is covered by CWC Schedule 2, but not by the AG Chemical Weapons Precursors list.

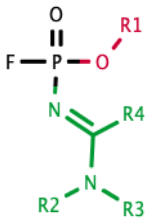
A

Amidine-featuring Novichok agents in CWC Schedule 1

Entry 1A13



Entry 1A14

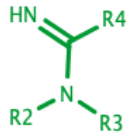


R1 and R4: H or alkyl (<=C10), including cycloalkyl

R2 and R3: Alkyl (<=C10), including cycloalkyl

B

Amidine Novichok Precursors in the AG Chemical Weapons Precursors list



	R2	R3	R4
AG 73	CH_3	CH_3	H
AG 74	CH_2CH_3	CH_2CH_3	H
AG 75	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$	H
AG 76	$CH(CH_3)_2$	$CH(CH_3)_2$	H
AG 77	CH_3	CH_3	CH_3
AG 78	CH_2CH_3	CH_2CH_3	CH_3
AG 79	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$	CH_3
AG 80	CH_3	CH_3	CH_2CH_3
AG 81	CH_2CH_3	CH_2CH_3	CH_2CH_3
AG 82	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$	CH_2CH_3
AG 83	CH_3	CH_3	$CH_2CH_2CH_3$
AG 84	CH_2CH_3	CH_2CH_3	$CH_2CH_2CH_3$
AG 85	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$
AG 86	$CH(CH_3)_2$	$CH(CH_3)_2$	$CH_2CH_2CH_3$
AG 87	CH_3	CH_3	$CH(CH_3)_2$
AG 88	CH_2CH_3	CH_2CH_3	$CH(CH_3)_2$
AG 89	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$	$CH(CH_3)_2$

Fig. 4. Panel A: CWC Schedule 1 covers two large families of Novichok agents with amidine branches. Panel B: the AG Chemical Weapons Precursors list covers only a limited subset of the amidines that can be used as precursors for the Novichoks listed in CWC Schedule 1.

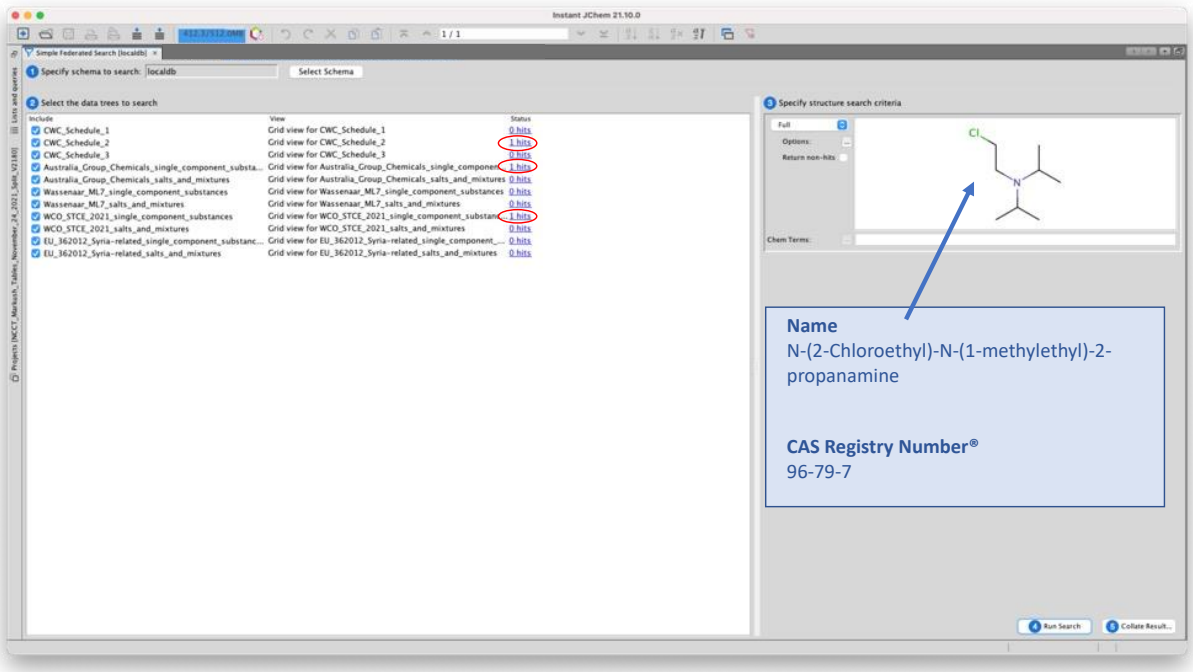
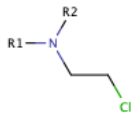
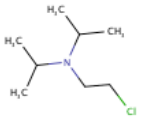


Fig. 5. The name and the CAS Registry Number of the chloroethylamine precursor for the synthesis of VX can be used as input to query the NCCT database. Instant JChem will convert the input into a structure. The structure is then checked against the lists implemented in the NCCT database. The search yields hits in CWC Schedule 2, the AG Chemical Weapons Precursors list, and the WCO STCE list.

CWC Schedule 2

Markush structure	Category	Entry Number	Entry Type	Entry Name	CAS Registry Number
	Precursors	CWC 2B10	Family	N,N-Dialkyl (Me, Et, n-Pr or i-Pr) aminoethyl-2-chlorides and corresponding protonated salts	---
	Overlaps	SMILES	InChI	InChIKey	PubChem URL
	AG 11, 54; STCE 19, 22, 79, 124, 125	---	---	---	

AG Chemical Weapons Precursors List

Markush structure	Category	Entry Number	Entry Type	Entry Name	CAS Registry Number
	Precursor	AG 11	Individual	N,N-Diisopropyl-(beta)-aminoethyl chloride	96-79-7
	Overlaps	SMILES	InChI	InChIKey	PubChem URL
	CWC 2B10 family member; STCE 19	CC(C)N(CCCl)C(C)C	InChI=1S/C8H18ClN/c1-7(2)10(6-5-9)8(3)4/h7-8H,5-6H2,1-4H3	InChIKey=DBVADBHSJCWFKI-UHFFFAOYSA-N	https://pubchem.ncbi.nlm.nih.gov/compound/7312

WCO STCE List

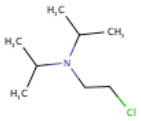
Markush structure	Category	Entry Number	Entry Type	Entry Name	CAS Registry Number
	Precursors	STCE 19	Individual	N,N-Diisopropyl-(beta)-aminoethyl chloride	96-79-7
	Overlaps	SMILES	InChI	InChI_Key	PubChem URL
	CWC 2B10 Family Member; AG 11	CC(C)N(CCCl)C(C)C	InChI=1S/C8H18ClN/c1-7(2)10(6-5-9)8(3)4/h7-8H,5-6H2,1-4H3	InChIKey=DBVADBHSJCWFKI-UHFFFAOYSA-N	https://pubchem.ncbi.nlm.nih.gov/compound/7312

Fig. 6. View of the hits yielded by the NCCT database search illustrated in **Fig. 5**. For individually listed chemicals, the NCCT database annotations includes structural identifiers (SMILES, InChi, and InChiKey) as well as a link to the PubChem entry relative to that specific chemical.

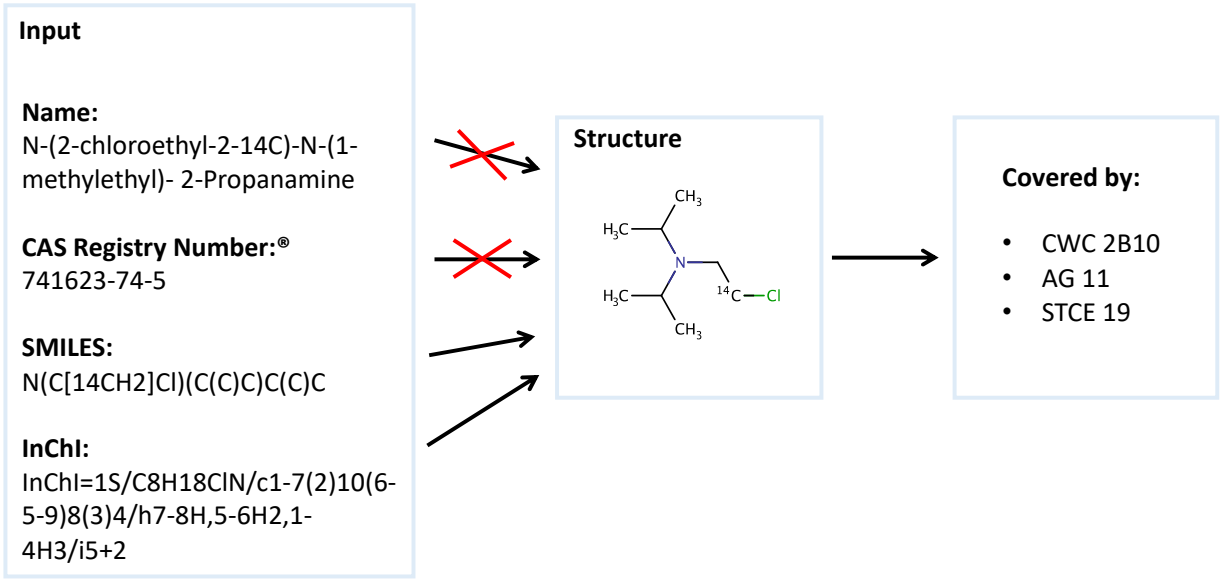


Fig. 7. The name and CAS Registry Number of an isotopically labeled version of the chloroethylamine precursor for the synthesis of VX cannot be converted into a structure by Instant JChem. Conversely, its structural identifiers are successfully converted into structures by Instant JChem. Just as in the case of the unlabeled variant, checking the structure against the NCCT database yields hits in CWC Schedule 2, the AG Chemical Weapons Precursors list, and the WCO STCE list.