molecular informatics

DOI: 10.1002/minf.200((full DOI will be filled in by the editorial staff))

Atom-to-atom mapping: a benchmarking study of popular mapping algorithms and consensus strategies

Arkadii Lin,[a] Natalia Dyubankova,[b] Timur I. Madzhidov,[c] Ramil I. Nugmanov,[c] Jonas Verhoeven,[b] Timur R. Gimadiev,[d] Valentina A. Afonina,[c] Zarina Ibragimova,[c] Asima Rakhimbekova,[c] Pavel Sidorov,[d] Andrei Gedich,[e] Rail Suleymanov,[e] Ravil Mukhametgaleev,[c] Joerg Wegner,[b] Hugo Ceulemans [b] and Alexandre Varnek*[a,d]

Abstract: In this paper, we compare the most popular Atom-to-Atom Mapping (AAM) tools: ChemAxon[1], Indigo[2], RDTool[3], NameRXN (NextMove)[4], and RXNMapper[5] which implement different AAM algorithms. An open-source RDTool program was optimized, and its modified version ("*new RDTool*") was considered together with several consensus mapping strategies. The Condensed Graph of Reaction approach was used to calculate chemical distances and develop the "AAM fixer" algorithm for an automatized correction of erroneous mapping. The benchmarking calculations were performed on a *Golden dataset* containing 1851 manually mapped and curated reactions. The best performing RXNMapper program together with the AMM Fixer was applied to map the USPTO database. The Golden dataset, mapped USPTO and optimized RDTool are available in the GitHub repository https://github.com/Laboratoire-deChemoinformatique.

Keywords: chemical reactions, data cleaning, atom-to-atom mapping

1 Introduction

Atom-to-atom mapping (AAM)[6,7] is a procedure that establishes a correspondence between the atoms of reactants and products. AAM allows to identify a reaction centre (RC) which, in turn, helps to prepare reaction templates used in an automatized forward/retrosynthesis planning[8–12], as well as to perform reaction classification[13] and reaction searching[14,15].

Several publicly and commercially available AAM tools are currently available. Their comparison would allow choosing an appropriate tool suitable for a given project. Some benchmarking studies have been already reported,[16,17] but they became outdated once new mappers were released.

In this paper, we compare the most popular AAM tools: Automapper (ChemAxon)[1], Indigo[2], RDTool[3], NameRXN (NextMove)[4], and RXNMapper[5] which implement different AAM algorithms. Thus, the open-source Indigo toolkit[2] applies the VF2 algorithm[18] which is an advanced graph matching algorithm based on Maximum Common Substructure (MCS) isomorphism principle[19]. This algorithm detects common structural motifs in reactants and products followed by identification of related pairs of atoms.

ChemAxon's Automapper uses a combination of MCS and a minimal Chemical Distance (CD) approaches. According to Chen et al.[20], a chemical distance is calculated as a sum of two terms: the number of chemical bonds created, broken, or transformed, and the number of atoms, which charge or valence state has been modified in chemical reaction. The AAM procedure with ChemAxon's Automapper consists of two steps: 1) detection of an appropriate MCS, and 2) CD-based optimization of mapping the atoms that are not a part of the MCS.

The NameRxn software[4] from NextMove performs AAM using large ensemble of rules describing known reaction mechanisms/transformations. The algorithm classifies a given reaction according to the implemented heuristics, followed by identification of the reaction centre described by an ensemble of atoms at chemical bonds formed, broken or transformed in course of chemical

[a] *Laboratory of Chemoinformatics, UMR 7140 CNRS, University of Strasbourg4, Blaise Pascal str., 67081 Strasbourg, France*
**e-mail: varnek@unistra.fr*

[b] *Janssen Pharmaceutica*
30, Turnhoutseweg str., 2340 Beerse, Belgium

[c] *Laboratory of Chemoinformatics and Molecular Modeling, Butlerov Institute of Chemistry, Kazan Federal University*
18, Kremlyovskaya str., 420008 Kazan, Russia

[d] *Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University*
Kita 21 Nishi 10, Kita-ku, 001-0021 Sapporo, Japan

[e] *Arcadia Inc.*
28 k2, Bolshoy Sampsonievskiy pr., St. Petersburg, 194044, Russia

Supporting Information for this article is available on the WWW under www.molinf.com

reaction. Once a reaction centre is defined, AAM becomes a simple procedure of simply numbering the atoms. Notice that no AAM is performed if the algorithm fails with reaction classification.

RXNMapper tool[5] implements a deep learning algorithm that is much faster than the abovementioned ones because it does not need to perform time-consuming combinatorial-seeking operations. RXNMapper uses a model based on the transformer architecture[21–23] which relates SMILES strings[24] of reactants (input) and products (output). An attention mechanism is used to assign a correspondence between the atoms of reactants and products. RXNMapper was trained in unsupervised settings on a subset of the USPTO database[25] for which the AAM accuracy >99% was reported.

Reaction Decoder Tool (or RDTool) uses a consensus approach combining four distinct MCS-based algorithms: *Mixture-MCS model*, *Min-sub model*, *Max-sub model*, and *Assimilation model*[26]. All of them are based on pairwise MCS identification between reactant and product molecules. In the *Min-sub model* (further mentioned as MIN), the smallest MCS between the reactants and products are first mapped. On the contrary, in the *Max-sub model* (further mentioned as MAX), the largest MCS between the reactants and products are first mapped. Then, the mapped part is removed from the reaction followed by the repetition of pairwise MCS identification until no more atoms can be mapped. The user-defined structural matches instead of MCS can be applied in preference. The *Mixture-MCS model* (further mentioned as MIXTURE) shares some features from MIN and MAX approaches. *Assimilation model* uses a restricted MCS algorithm that prefers the ring-to-ring mapping over the ring-to-chain one. Among those, the best AAM solution is selected according to a score accounting for (*i*) fragments count, (*ii*) dynamic bonds count, and (*iii*) energy of bonds cleavage/formation. According to Rahman et al.[3], the best solution corresponds to the minimal structural changes. In line with this principle, the algorithm minimizes the bond changes in the reaction and, therefore, reduces the chemical distance and energy costs for forming or breaking the bonds. To compute the energy of bonds cleavage/formation, the precomputed values are used [27–29]. Notice that the number of dynamic bonds is counted using the Dugundji-Ugi matrix[30], which requires a fully balanced reaction equation. For this reason, RDTool performs some data pre-processing including reaction balancing. This, unfortunately, cancels previously performed reaction standardization.

It should be noted that any mapper assigns to the atoms the numbers (hereafter, called “AAM numbers”) further used in the reaction SMILES or RXN file to identify atom-to-atom mapping. Among the benchmarked mappers, only ChemAxon mapper uses the initial ordering of atoms as AAM numbers.

All benchmarking calculations were performed on a reference *Golden* dataset containing 1851 manually mapped and curated reactions.

2 Method

2.1 Optimized RDTool and consensus solutions

Since RDTool is an open-source software, we decided to refactor the code in order to use independently three out of four constituent algorithms: Min-sub (*MIN*), Max-sub (*MAX*) and Mixture-MCS (*MIXTURE*) models. The fourth algorithm was omitted due to the code fuzziness. The code was also revised in order to accelerate the functioning of the program. This new version of the program was named *new RDTool*.

In complement to individual AAM tools, six consensus scenarios were considered:

1) *Consensus 1* - RXNMapper & MAX & MIXTURE & MIN & ChemAxon & Indigo & NameRXN;
2) *Consensus 2* - RXNMapper & MAX & MIXTURE & MIN & ChemAxon & Indigo;
3) *Consensus 3* - RXNMapper & MAX & MIXTURE & MIN & ChemAxon;
4) *Consensus 4* - RXNMapper & MAX & ChemAxon;
5) *Consensus 5* - RXNMapper & ChemAxon;
6) *Consensus 6* – MAX & MIN & MIXTURE.

Within these six consensus strategies, each individual/tool generates its AAM solution, followed by selection of the best one according to minimal Chemical Distance score.

2.2 Assessment of AAM accuracy with Condensed Graph of Reaction

Chemical Distance corresponding to a given AAM solution can be easily assessed for any reaction transformed into a Condensed Graph of a Reaction (CGR)[31,32]. CGR is a single graph, which encodes an ensemble of reactants and products as shown in Figure 1. CGR results from the superposition of the atoms of products and reactants having the same numbers. The nodes in CGR correspond to “static” or “dynamic” atoms whereas the edges correspond to “static” or “dynamic” bonds. The state of dynamic objects - charge and radical label for atoms and bond order for bonds - is changed in course of reaction. Once the AAM operation is performed, a related CGR can be prepared even if a reaction equation is not balanced (e.g., a minor reactant or a leaving group is missed). In line with the Chen et al.[20] definition, a Chemical Distance for a given reaction is estimated by the number of related dynamic objects. Clearly, different AAM solutions lead to different CGRs.

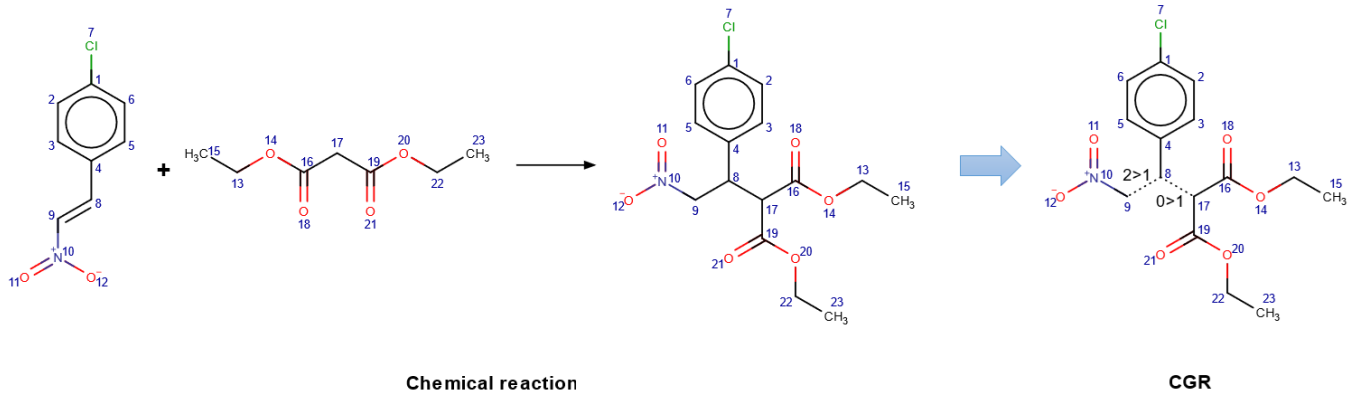

Figure 1. The reaction between malonic ester and 4-chloro-beta-nitrostyrene (left) and related Condensed Graph of a Reaction (CGR) (right). Dynamic bonds in CGR are depicted by dash lines. The dynamic bond "2>1" between two carbon atoms 8 and 9 corresponds to the transformation of a double bond to a single bond, whereas "0>1" corresponds to the formed single bond between atoms 8 and 17. The reaction centre (RC) includes atoms 8, 9, and 17 at the dynamic bonds.

In this study, CGRs have been prepared for all reactions from the reference set according to AAM performed either manually by experts in organic chemistry or using considered software tools. Ideally, a mapping provided by a particular tool corresponds to the manual one, i.e., the related CGRs should be identical. Notice that in some cases, erroneous mapping leads to the same chemical distance as that issued from the manual AAM (see example in Section 2.3). Therefore, the identity of two condensed graphs is a more solid criterion of the AAM correctness than a simple comparison of related chemical distances.

2.3 Heuristics for correction of erroneous AAM

Systematic errors of atom-to-atom mapping can be corrected using heuristics based on the CGR approach. The idea is to replace a structural motif representing an erroneous reaction centre and its environment in a CGR with another one corresponding to a correct mapping using empirical remapping rules. The concept of AAM fixing was proposed in our early publication[33]; here we demonstrate its fully automatized implementation.

The overall process can be split into two parts: *(i)* fixing rules preparation (see Figure 2), and *(ii)* fixing rules application (see Figure 3). A fixing rule is generated for a pair of correct/erroneous AAM generated for a given reaction and represented in either RXN or SMILES formats. A fixing rule is generated in 4 steps:

1. Two related CGRs are prepared for correctly and erroneously mapped reactions. For each CGR, a reaction centre with its closest environment (RC_{close}) is extracted (see Figure 2a).
2. Preparation of "extended reaction centres", RC_{ext}, by augmentation of correct RC_{close} fragment by the atoms belonging to the erroneous RC_{close} and vice versa. In such a way, both correct and erroneous RC_{ext} fragments contain the same set of atoms (Figure 2b).
3. In order to determine the required permutations ***P*** of AAM numbers, the extended reaction centres are deconvoluted to reaction equations. A permutation rule can be obtained by overlapping the atomic AAM numbers' lists extracted from the corresponding deconvoluted reaction equations (Figure 2c).
4. A fixing rule is compiled out of three instances: an erroneous and correct RC_{ext} and a permutation rule (Figure 2d).

Notice that the order of atoms in the two atoms' lists at step 3 must be the same: the algorithm crashes or generates wrong permutation solutions when the atoms are shuffled by the mapper. Among all benchmarked here tools, only ChemAxon mapper keeps the order of atoms as in the initial RXN of SMILES. Therefore, AAM Fixer was applied only to the reactions mapped by ChemAxon.

Once a fixing rule is prepared, it can be applied to fix AAM numbers for a given reaction for which a CGR is generated (Figure 3a). Then erroneous extended reaction centre is embedded in the CGR (Figure 3b). Hits found for a particular RC_{ext} signal that the AAM numbers are wrong and can be fixed with the corresponding permutation rule ***P*** (Figure 3c). AAM numbers in the product part of the reaction are permuted according to ***P***. Finally, a new CGR is generated for the corrected reaction followed by the embedding of the correct RC_{ext} in order to check whether the AAM fixing was done successfully (Figure 3d). The original AAM numbers are restored if the latter embedding fails. Once extracted, the fixing rules can be applied to correct AAM generated by any mappers.

Fixing rule extracted from a given reaction is validated by its application to the same reaction. In some cases, the correct mapping cannot be achieved. Usually, it concerns the cases when atoms distant from the reaction center need to be remapped, see Figure S1 in SI. Finally, all validated fixing rules are integrated into AAM Fixer.

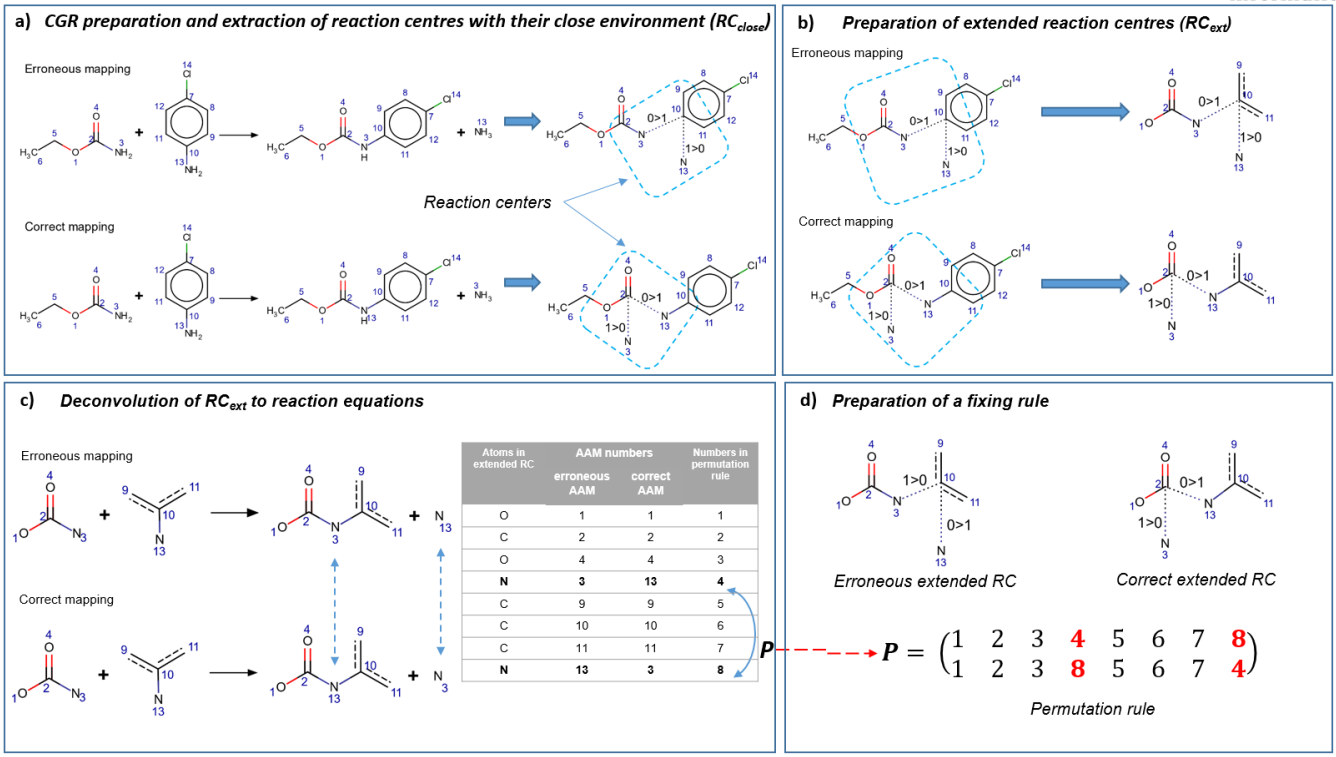

Figure 2. AAM fixing rule preparation workflow illustrated on the example of the addition-elimination reaction for which erroneous mapping was obtained with ChemAxon. (a) The RC_{close} fragments corresponding to the correct and erroneous mapping include six atoms only four of which (atoms 2, 3, 10 and 13) are in common. (b). Eight atoms, that belong to correct or wrong reaction centres or shared by both (atoms 1, 2, 3, 4, 9, 10, 11 and 13) are considered in RC_{ext}. Thus, extended reaction centres are prepared by the addition of atoms 9 and 11 to the correct RC_{close} fragment and of atoms 1 and 4 to the erroneous RC_{close}. (c) Deconvolution of RC_{ext} to reaction equations and superposition of the related atoms in the products reveals a permutation rule. (d) The numbers of the atoms to be permuted are shown in red. The dynamic bond "1>0" in CGRs corresponds to a broken single bond, and "0>1" corresponds to a formed single bond. Notice, that the chemical distances (the number of dynamic bonds in CGR) are equal to 2 for both erroneously and correctly mapped reactions.

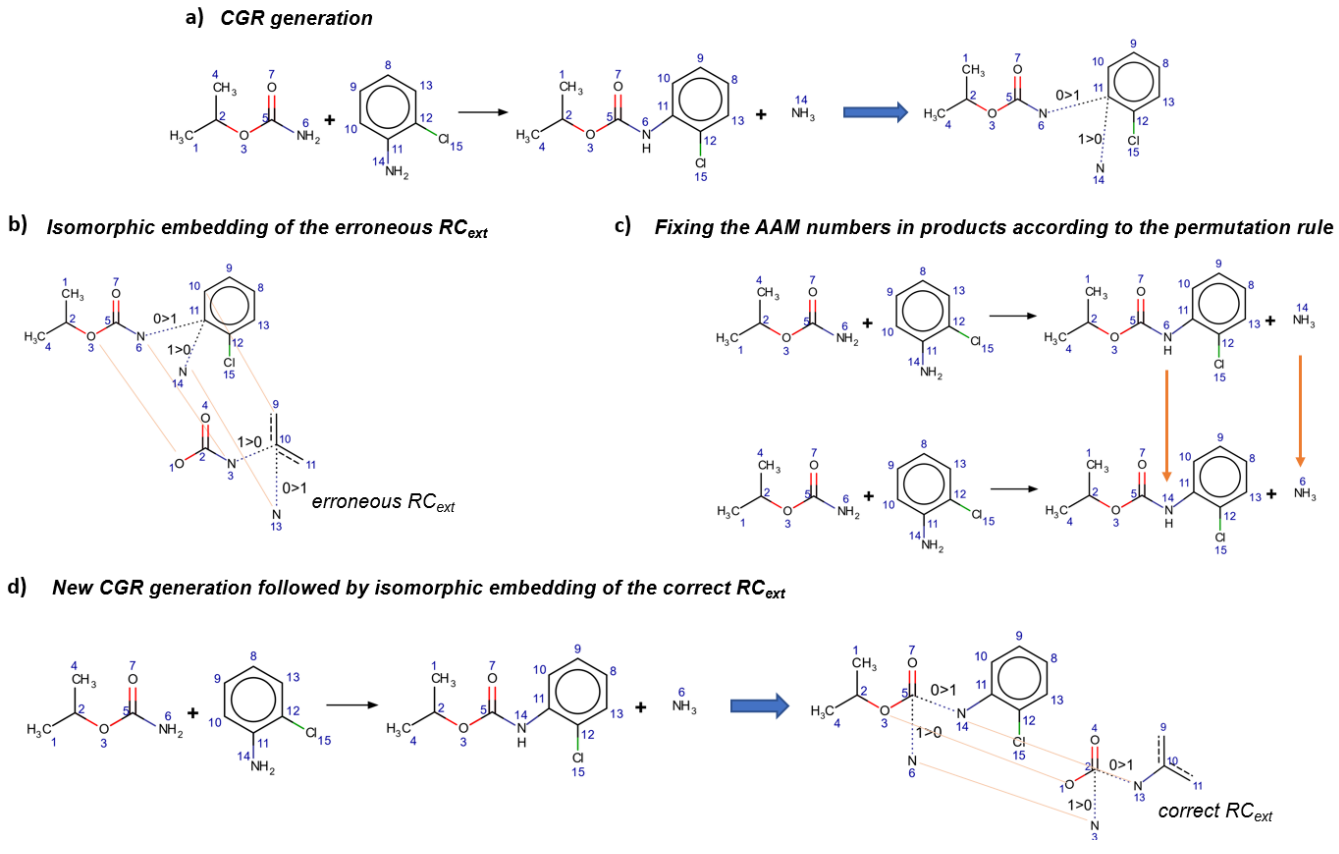

Figure 3. Example of a fixing rule application. Here, an addition-elimination reaction was mapped by ChemAxon toolkit which mixed up the nitrogen atoms N_6 and N_{14}. To fix this problem, (*a*) a CGR is generated, and (*b*) isomorphic embedding of an erroneous RC_{ext} is performed. *(c)* AAM atomic labels of the nitrogens N_6 and N_{14} in the product are permutated according to the fixing rule. Finally, *(d)* a new CGR is generated, and the correct RC_{ext} is embedded to check if the AAM correction is done successfully. Here, the dynamic bond "1>0" corresponds to a broken single bond, and "0>1" corresponds to a formed single bond.

2.4 Data

The data set used in this study as a reference resulted from the merging of curated dataset used in the benchmarking study by Jaworski et al.[34] and a subset of reactions extracted from USPTO database. The first set contains reactions with full stoichiometry and taken from the Organic Syntheses collection[35], randomly selected and typically mostly stoichiometrically unbalanced reactions from Reaxys[36], mechanistically complex reactions taken from various literature sources (e.g., Kurti's "Strategic Application of Named Reactions in Organic Synthesis"[37] or Grossman's "The Art of Writing Reasonable Organic Reaction Mechanisms"[38]), relatively simple reactions with full stoichiometry taken from total syntheses published in Org. Lett., J. Am. Chem. Soc., and J. Org. Chem., typical reactions without full stoichiometry extracted from patents, and mechanistically complex reactions which include rearrangements and multicomponent reactions taken from recent literature (in most cases, after 2010 and from Org. Lett., J. Am. Chem. Soc., and J. Org. Chem.). A detailed description of reactions types included into the dataset can be found in ref.[34] The second set extracted from USPTO gathers examples of the most popular reaction types used in medicinal chemistry (e.g. S_N2, Buchwald-Hartwig, Suzuki-Miyaura coupling, etc.)[39].

A manual check of 1405 organic reactions from the Jaworski's set[34] revealed several serious problems with structures preparation or atom-to-atom mapping. Thus, 10 reactions were identified as duplicated (see an example in Figure S2 in Supporting Information), 30 reactions were discarded since they contained molecules with invalid valences (Figure S3) or they could not be interpreted (Figure S4). The discarded reactions were analyzed manually and curated, if possible. Namely, 17 reactions with chemical structure representation errors were corrected. We've also discovered some 350 reactions containing radicals that were manually curated (see an example of such reaction in Figure S5).

Manual inspection of atom-to-atom also revealed several problems. Eight reactions containing obvious errors (*e.g.,* one carbon atom in a reactant corresponded to both oxygen and carbon atoms in a product simultaneously) were manually curated (see an example in Figure S6). On the other hand, 22 reactions with non-obvious mapping which may depend on experimental conditions (see Figure S7), and 3 reactions with unclear reaction mechanisms or lost reactants (see Figure S8) were removed.

Curated Jaworski's dataset containing 1382 reactions was extended with 469 manually mapped USPTO reactions. The final reference ("*Golden"*) dataset comprises 1851 reactions. Its composition has been analysed with the help of the NameRxn (NextMove) tool. The tool failed to classify 628 reactions. Occurrences of particular reaction types for remaining 1164 reactions are given in Supporting Information (see Figure S9 in SI).

In addition, the entire USPTO database[25] containing some 1.3M unique and standardized reactions[40] (functional groups standardization, dearomatization, valence checking, selection of major tautomers, removal of explicit hydrogens, cleaning radicals and isotopes, clear stereo, split of ions, aromatization, duplicates removal) was used in order to demonstrate the functioning of AAM Fixer.

3 Results and Discussion

3.1 Benchmarking results

15 mapping strategies described in the Method section were applied to the *Golden* dataset. As it follows from Figure 4, ChemAxon is the fastest tool: it takes approximately 0.003 second to map one reaction. NameRXN (NextMove) and RXNMapper follow ChemAxon with 0.02 and 0.05 seconds per reaction, respectively. RDTool is the slowest one: it takes around 2.46 seconds to map one reaction. This is not surprizing since RDTool executes four mapping algorithms followed by calculation of seven metrics used to perform consensus prediction. Revision of its code significantly accelerated the mapping procedure: the new RDTool including three AAM algorithms (MIN, MAX, and MIXTURE) is about 40% faster than the original tool.

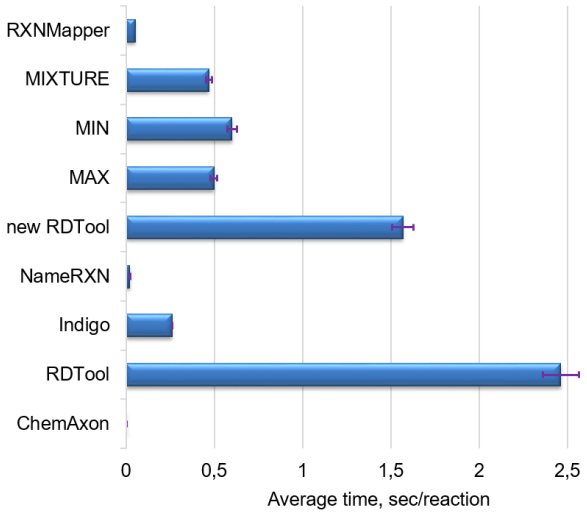

Figure 4. Average CPU time per reaction for different AAM tools .

Figure 5 and Table 1 report an information about AAM performance of different tools and algorithms assesed for the Golden set. As explained in Section 2.2, AAM performed for a given reaction was considered correct if related CGR coincided with that obtained by the manual mapping.

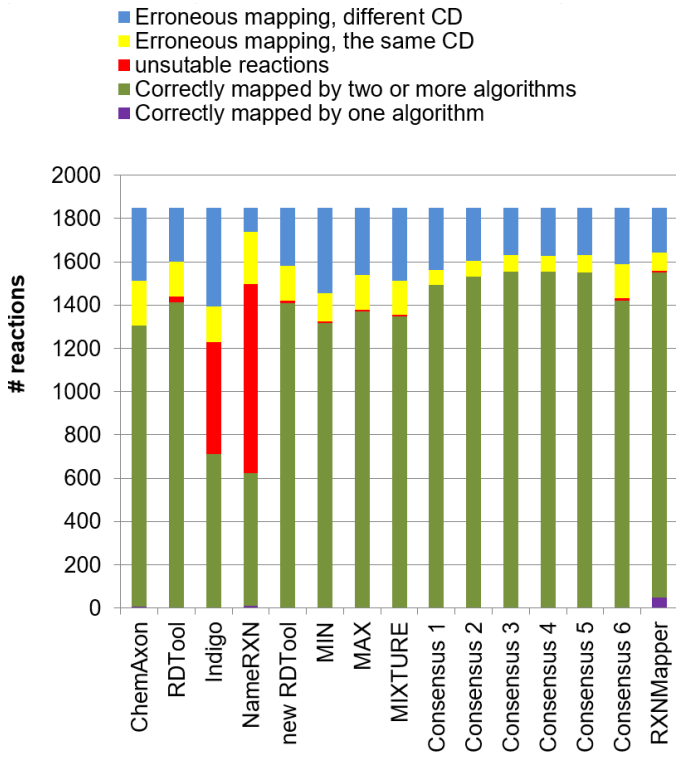

Figure 5. Performance of different AAM tools and consensus strategies.

Compared to other tools, NameRXN[41] demonstrates the lowest rate of correct atom-to-atom mapping (33.6%). This can be explained by the fact that the AAM algorithm uses a limited set of reaction classification rules. If there is no rule characterizing a given reaction, no mapping is performed. In addition, existing rules do not always guarantee a correct AAM (see Figure 6).

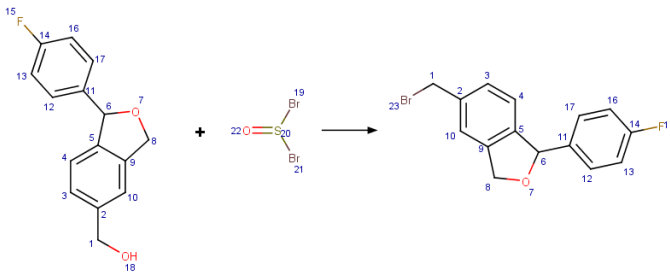

Figure 6. Example of a bromination reaction wrongly mapped by the NameRXN tool. In this reaction, one of the bromine atoms of thionyl bromide (number 19 or 21) replaces the OH group in [1-(4-fluorophenyl)-1,3-dihydro-2-benzofuran-5-yl]methanol. However, in the product, the tool assigns number 23 to the Br atom, which formally corresponds to a fictive reactant missed in the reaction equation.

The RXNMapper tool demonstrates a good performance: 1550 reactions (83.74%) were correctly mapped which is the best result among all considered reaction mappers. Unlike other studied mappers, RXNMapper consistently generates correct AAM for esterification and esters hydrolysis reactions. However, this tool has a serious drawback: it resets any previously performed standardization operation and reorders the reactants and the products (Figure S10).

The original RDTool succeeded to map 76.23% of reactions in the reference data set, which is very similar to the new RDTool (76.18%). The code revision performed in this work allowed us to significantly reduce the time of computations (see Figure 4) and to apply MIN, MAX, and MIXTURE algorithms separately (see Method section). Comparing performances of the new RDTool and its constituting individual mapping algorithms, we found that the MAX algorithm already provides a 73.96% accuracy, which corresponds to 97.1% of the total number of reactions correctly mapped by the new RDTool. Thus, the MAX algorithm is almost as accurate as the entire RDTool but is 5 times faster.

The accuracy of ChemAxon's AAM (70.45%) is similar to earlier reported results.[5,34] Overall, 547 reactions were erroneously mapped.

The success rate of the Indigo mapper is relatively low, especially for halogenation reactions. In some cases, the tool maps a halogen atom in a product as it comes from an unspecified reactant, even if a Hal_2 molecule is present in the reactants part (see an example in Figure S11 in SI). In addition, in the case of symmetric molecules, Indigo makes one-to-many correspondence of one reactant atom to several product atoms, i.e., it assigns the same numbers to the atoms in different symmetrically identical fragments (see an example in Figure S12). This leads to significant problems with further usage of mapping results, e.g., with CGR generation. Therefore, reactions with one-to-many correspondence were classified as "unsuitable" (see Figure 5 in red).

Among all examined consensus solutions, only *Consensus 3* combining RXNMapper, MAX, MIXTURE, MIN and ChemAxon methods performs slightly better than RXNMapper individually (84.06% vs 83.74%). This negligible accuracy improvement does not compensate a 30 folds slowdown of the consensus compared to the individual solution.

Besides that, Table 1 demonstrates that the lowest Chemical Distance (CD) does not always guarantee correct AAM assignment. Thus, for a large portion of reactions, erroneous and correct mapping correspond to equal chemical distances whatever the mapping algorithm is. Moreover, in some Diels-Alder reactions, the lowest CD corresponds to erroneous AAM (see Figure 7).

Another question concerns the difference in the AAM algorithms implemented in the benchmarked tools. Table 1 shows that the number of reactions correctly mapped with only one selected algorithm is negligible except of that for RXNMapper which is the only tool performing correct mapping for esterification and esters hydrolysis reactions on the Golden dataset. The latter, however, wasn't reproduced for the entire set of USPTO reactions. It is also surprising that according to our tests, the tool perfectly maps reactions of some particular types in the *Golden* dataset, but it fails on those from USPTO[25]. In particular, this concerns reduction of carboxyl groups to alcohols (89% reactions of this type in USPTO were erroneously mapped), methyl ester formation in methanol (31%), and esterification reactions (10-30%).

Table 1. AAM benchmarking results.

AAM tool or consensus strategy	# correctly mapped reactions	Accuracy,%	#reactions correctly mapped by the given tool only	# "unsuitable" reactions [a]	# erroneous AAM with different CD[b]	# erroneous AAM with the same CD[c]
ChemAxon	1304	70.45	5	0	339	208
Indigo	712	38.47	4	518	457	164
NameRXN	622	33.60	9	873	113	243
RXNMapper	1550	83.74	49	9	208	84
RDTool	1411	76.23	0	30	250	160
New RDTool	1410	76.18	0	9	269	163
MIN	1316	71.10	0	8	395	132
MAX	1369	73.96	0	8	311	163
MIXTURE	1347	72.77	0	8	339	157
Consensus 1[d]	1494	80.71	-[e]	0	289	68
Consensus 2	1532	82.77	-	0	246	73
Consensus 3	1556	84.06	-	0	220	75
Consensus 4	1554	83.95	-	0	222	75
Consensus 5	1549	83.68	-	0	221	81
Consensus 6	1422	76.82	0	8	262	159

[a] The mapping tool can not map a given reaction

[b] Chemical Distance for the erroneously mapped reaction differs from that with correct AAM.

[c] Chemical Distance for the erroneously mapped reaction is equal to that with correct AAM.

[d] See description of the consensus strategies in Section 2.1.

[e] not relevant for consensus strategies

3.2 Correction of erroneous mapping with the AAM Fixer

In order to investigate the efficiency of the AAM Fixer tool, it has been applied to the mapped Golden dataset aiming (i) to extract the fixing rules and (ii) to correct erroneously mapped reactions.

For 547 (out of 1851) reactions incorrectly mapped by ChemAxon, 535 unique fixing rules were extracted. Out of those, only 442 rules have been validated (see section 2.3) which allowed us to correct 452 reactions with erroneous mappings. Interestingly, that the rules derived for ChemAxon mapping allowed us to correct 62 reactions from the Golden set erroneously mapped by RXNMapper. This proves the transferability of the fixing rules. Unfortunately, it is impossible for the moment to extract fixing rules for reactions mapped by RXNMapper since all considered here AAM algorithms except ChemAxon mapper change the atoms' order in the RXN file (see Section 2.3).

Finally, AAM Fixer powered by the rules, extracted for the Golden dataset, was applied to the USPTO dataset mapped by RXNmapper. This led to AAM corrections for 2021 reactions. Such a small portion of fixed reactions is caused by the fact that rules are quite selective due to the RC environment of radius 1 (RC_{close}).

The advantage of the AAM Fixer is its ability to fix AAM in cases when the chemical distance for a wrong AAM is the same or even lower than for a correct one, like Diels-Alder or esterification reactions (Figure S13 in Supporting Materials). Notice that the fixing rules can be prepared manually and then implemented in the AAM fixer.

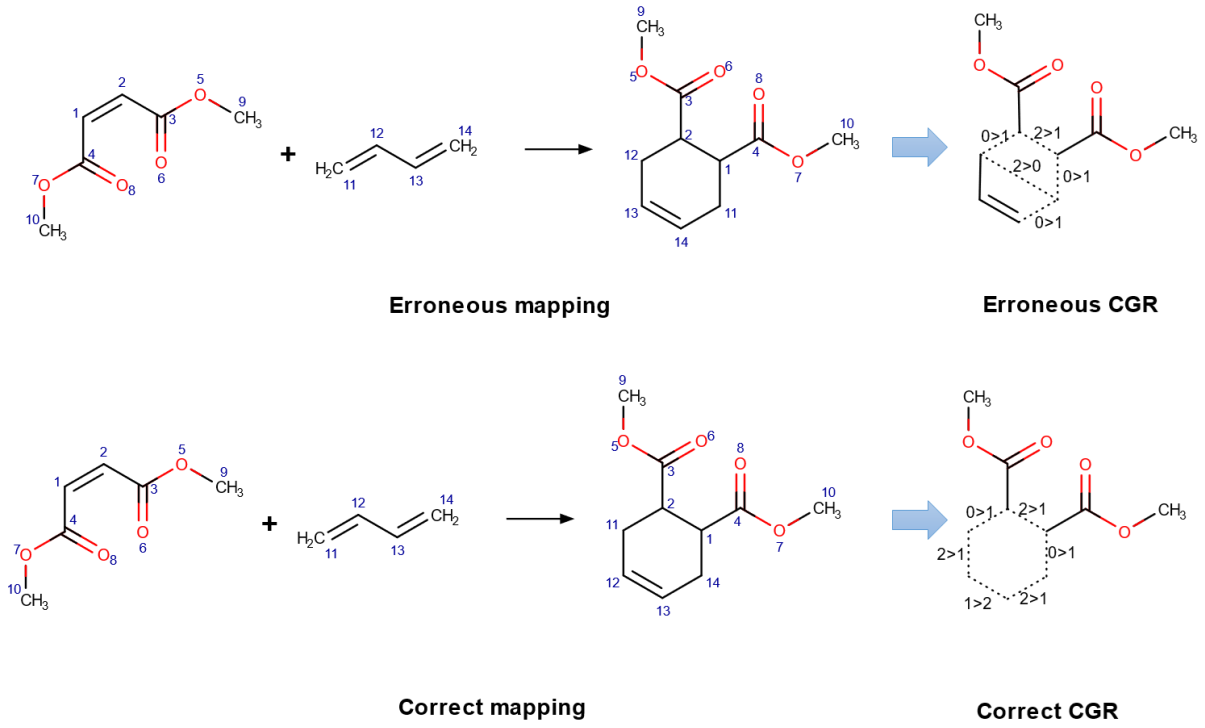

Figure 7. Example of Diels-Alder reaction with erroneous (top) and correct (bottom) atom-to-atom mappings. Here, the reaction itself is on the left side, and the corresponding CGR is on the right side. The symbol "*n*>*m*" corresponds to the transformation of bond order *n* to *m*[33]. Notice that the chemical distance corresponding to erroneous mapping (CD=5) is lower than that for correctly mapped reaction (CD=6).

4 Conclusions

Benchmarking study comparing the performance of Atom-to-Atom Mapping of 5 popular tools (RXNMapper, ChemAxon, NameRXN, Indigo, and RDTools) and several algorithms and consensus strategies applied to the "Golden" dataset revealed that no one provides with an ideal AAM assignment. RXNMapper is ranked the best with the accuracy 83.74% and a reasonable speed of calculations (0.05 seconds). However, the deep learning model integrated into this tool looks overfitted because it fails to map certain types of USPTO reactions although very similar reactions from the Golden set were perfectly mapped.

It has been demonstrated that the lowest Chemical Distance doesn't always guarantee correct AAM assignment. Thus, for some Diels-Alder reactions, erroneous mapping corresponds to smaller CDs compared to correct AAM. For a large portion of reactions, erroneous and correct mapping correspond to equal chemical distances for any studied tool. For such cases, we have developed a rule-based AAM Fixer tool that corrects frequent types of AAM errors. It extracts rules from a pair of correct and wrong AAM and can fix mapping for new reactions.

Finally, we've proven that the Condensed Graph of Reaction approach is particularly useful for chemical distances assessment, detection and correction of wrong mapping.

5 Data availability.

The manually mapped "Golden" dataset containing 1851 reactions, the USPTO database mapped with RXNMapper, standardization workflow and the optimized RDTool (*new RDTool*) are available on GitHub: https://github.com/Laboratoire-de-Chemoinformatique/.The AAM Fixer tool with embedded rules is available on GitHub: https://github.com/cimm-kzn/AAMFixer .

6 Funding

This project was supported by the VLAIO HBC.2018.2287 MaDeSMArt (Machine Design of Small Molecules by AI) grant. The AAMFixer approach development was supported by the subsidy allocated to Kazan Federal University for the state assignment in the sphere of scientific activities.

7 Conflict of interests.

There is no conflict of interest

8 Author contribution statement. The authors equally participated in this work.

9 Abbreviations

AAM - atom-to-atom mapping; CGR - Condensed Graph of Reaction; CD - Chemical Distance.

10 References

[1] "AutoMapper," can be found under https://docs.chemaxon.com/display/docs/automapper-user-s-guide.md, **2020**.

[2] A. Savelev, I. Puzanov, V. Samoilov, V. Karnaukhov, **2019**.

[3] S. A. Rahman, G. Torrance, L. Baldacci, S. Martínez Cuesta, F. Fenninger, N. Gopal, S. Choudhary, J. W. May, G. L. Holliday, C. Steinbeck, J. M. Thornton, *Bioinformatics* **2016**, *32*, 2065–2066.

[4] NextMove Software, "NameRXN," can be found under www.nextmovesoftware.com, **2020**.

[5] P. Schwaller, B. Hoover, J. L. Reymond, H. Strobelt, T. Laino, *ChemRxiv* **2020**, DOI 10.26434/chemrxiv.12298559.v1.

[6] E. Fontain, *Anal. Chim. Acta* **1992**, *265*, 227–232.

[7] W. L. Chen, D. Z. Chen, K. T. Taylor, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 560–593.

[8] J. Law, Z. Zsoldos, A. Simon, D. Reid, Y. Liu, S. Y. Khew, A. P. Johnson, S. Major, R. A. Wade, H. Y. Ando, *J. Chem. Inf. Model.* **2009**, *49*, 593–602.

[9] A. Bøgevig, H. J. Federsel, F. Huerta, M. G. Hutchings, H. Kraut, T. Langer, P. Löw, C. Oppawsky, T. Rein, H. Saller, *Org. Process Res. Dev.* **2015**, *19*, 357–368.

[10] M. H. S. Segler, M. P. Waller, *Chem. – A Eur. J.* **2017**, *23*, 5966–5971.

[11] M. E. Fortunato, C. W. Coley, B. C. Barnes, K. F. Jensen, *ChemRxiv* **2020**, DOI 10.26434/chemrxiv.11811564.v1.

[12] C. W. Coley, W. H. Green, K. F. Jensen, *Acc. Chem. Res.* **2018**, *51*, 1281–1289.

[13] J. B. Hendrickson, L. Chen, in *Encycl. Comput. Chem.* (Ed.: P. von Ragué Schleyer), John Wiley & Sons, Ltd, Erlangen, **2002**, pp. 2381–2402.

[14] A. I. Lin, T. I. Madzhidov, O. Klimchuk, R. I. Nugmanov, I. S. Antipin, A. Varnek, *J. Chem. Inf. Model.* **2016**, *56*, 2140–2148.

[15] L. Chen, J. G. Nourse, B. D. Christie, B. A. Leland, D. L. Grier, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1296–1310.

[16] E. E. Litsa, M. I. Peña, M. Moll, G. Giannakopoulos, G. N. Bennett, L. E. Kavraki, *J. Chem. Inf. Model.* **2019**, *59*, 1121–1135.

[17] G. A. Preciat Gonzalez, L. R. P. El Assal, A. Noronha, I. Thiele, H. S. Haraldsdóttir, R. M. T. Fleming, *J. Cheminform.* **2017**, *9*, 39.

[18] L. P. Cordella, P. Foggia, C. Sansone, M. Vento, *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 1367–1372.

[19] J. W. Raymond, P. Willett, *J. Comput. Aided. Mol. Des.* **2002**, *16*, 521–533.

[20] L. Chen, J. Gasteiger, J. R. Rose, *J. Org. Chem.* **1995**, *60*, 8002–8014.

[21] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, in *ArXiv* (Eds.: I. Guyon, U. V Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett), Curran Associates, Inc., **2017**, pp. 5998–6008.

[22] J. Devlin, M. W. Chang, K. Lee, K. Toutanova, *arXiv* **2018**.

[23] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, R. Soricut, *arXiv* **2019**.

[24] D. Weininger, *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.

[25] D. Lowe, **2017**, DOI 10.6084/m9.figshare.5104873.v1.

[26] S. A. Rahman, S. M. Cuesta, N. Furnham, G. L. Holliday, J. M. Thornton, *Nat. Methods* **2014**, *11*, 171–174.

[27] D. P. Stevenson, *The Strengths of Chemical Bonds*, Butterworths Scientific Publications, **1955**.

[28] B. deB, **1970**.

[29] S. W. Benson, *J. Chem. Educ.* **1965**, *42*, 502.

[30] J. Dugundji, I. Ugi, in *Comput. Chem.*, Springer Berlin Heidelberg, Berlin, Heidelberg, **2006**, pp. 19–64.

[31] F. Hoonakker, N. Lachiche, A. Varnek, *Int. J. Artif. Intell. Tools* **2011**, *20*, 253–270.

[32] T. I. Madzhidov, P. G. Polishchuk, R. I. Nugmanov, A. V. Bodrov, A. I. Lin, I. I. Baskin, A. A. Varnek, I. S. Antipin, *Russ. J. Org. Chem.* **2014**, *50*, 459–463.

[33] R. I. Nugmanov, R. N. Mukhametgaleev, T. Akhmetshin, T. R. Gimadiev, V. A. Afonina, T. I. Madzhidov, A. Varnek, *J. Chem. Inf. Model.* **2019**, *59*, 2516–2521.

[34] W. Jaworski, S. Szymkuć, B. Mikulak-Klucznik, K. Piecuch, T. Klucznik, M. Kaźmierowski, J. Rydzewski, A. Gambin, B. A. Grzybowski, *Nat. Commun.* **2019**, *10*, 1434.

[35] "Organic Syntheses," can be found under http://www.orgsyn.org/, **2021**.

[36] J. Goodman, *J. Chem. Inf. Model.* **2009**, *49*, 2897–2898.

[37] L. Kurti, B. Czakó, *Strategic Applications of Named Reactions in Organic Synthesis*, Elsevier, **2005**.

[38] R. B. Grossman, R. Grossman, *The Art of Writing Reasonable Organic Reaction Mechanisms*, Springer, **2003**.

[39] D. G. Brown, J. Boström, *J. Med. Chem.* **2016**, *59*, 4443–4458.

[40] A. Lin, N. Dyubankova, T. Madzhidov, R. Nugmanov, A. Rakhimbekova, Z. Ibragimova, T. Akhmetshin, T. R. Gimadiev, R. Suleymanov, J. Verhoeven, J. K. Wegner, H. Ceulemans, A. Varnek, *ChemRxiv* **2020**, 1–20.

[41] N. Schneider, D. M. Lowe, R. A. Sayle, M. A. Tarselli, G. A. Landrum, *J. Med. Chem.* **2016**, *59*, 4385–4402.

Received: ((will be filled in by the editorial staff))
Accepted: ((will be filled in by the editorial staff))
Published online: ((will be filled in by the editorial staff))