ChemDiv

TACKLING THE PROBLEM OF HIT RATE ENRICHMENT BY VIRTUAL SCREENING: QUALITY-DRIVEN BIOACTIVE CONFORMATIONS AND PHARMACOPHORE-BIASED CREATIVE SEARCH

Nasonov, Anatoliy F. , ChemDiv, Inc, 6605 Nancy Ridge Drive, San Diego, CA 92121 USA;
Pletnev Igor V., Department of Chemistry, Moscow State University, 119992 Moscow, Russia

Introduction

ChemoSoft™ is an out-of-the box advanced software solution for drug-design, combinatorial and classical chemistry[1]. Recently we developed ChemoSoft™'s 3D Tools processing 3D molecular structures. The problem of generation of bioactive conformations have been attracting continuous interest [2,3]. There are controversial opinions on what approach gives better results: diversity-oriented, rule-based or low-conformation-directed [2, 3]

To investigate this issue, we designed the following advanced features of our tools:
I. 2D-3D Converter joint with conformational analysis which
-generates an ensemble of conformers prior to geometry optimization
-selects several low energy candidates from the ensemble for geometry optimization
-chooses the energy champion from the optimized candidates
-employs full molecular mechanics
-optimizes not only rings but also chains
-adjusts not only torsional angles but also valent angles and bond lengths during geometry optimization.

II. 3D pharmacophore search tool allowing creative search - pharmacophore query-biased search with
-ability to generate multiple conformers from 2D input coordinates of structures
-ability to filter results of systematic conformational search at the first stages of 2D-3D conversion of molecular structures relative to a preliminary set pharmacophore query (based on a bioactive conformation). In our study we compare our software results with those published earlier [2, 3].

Computational methods

Hardware and software environment

All computations were performed on a computer with AMD® Turion™ 64 ML-30 1.6 GHz processor in WindowsXP™ operating system.

The superimposition procedure

All conformations were superimposed by using a command line utility called Strfit (version 2.0 designed by I. Pletnev).

Molecular modeling and pharmacophore search

All molecular modeling and pharmacophore search procedures were performed with the use of ChemoSoft's 3D Tools™ version 2.70 (see below).

2D-3D Conversion

All studied structures were converted from 2D connection table information obtained from unmodified X-ray ligand structures saved as MDL® mol files by manually assigning double and aromatic bonds to the extracted X-ray structures according to reference [2] and automatic using 2D connectivity information from the 3D saved files. The latter was done by ChemoSoft's 2D-3D Converter™ which is a part of ChemoSoft's 3D Tools™ and consists of a command line utility Xcgen and a GUI to it, called Converter. We used the mode with input bonding and stereochemistry information retained (if available). The number of output conformers was set to 1. For visual inspection convenience we also did not add hydrogens during 2D-3D conversion. A typical GUI window of ChemoSoft's 2D-3D Converter™ is shown in Fig 1.

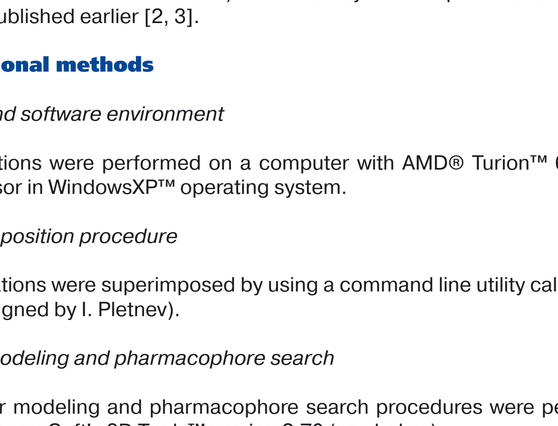


Figure 1. The window of ChemoSoft's 2D-3D Converter™.

The obtained conformations were visually inspected by ChemoSoft's 3D Viewer™ (which is a part of ChemoSoft's 3D Tools™) as shown in Fig 2.

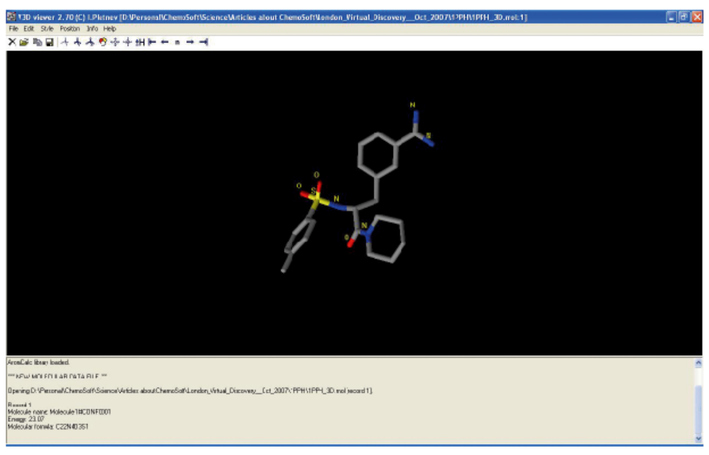


Figure 2. The window of ChemoSoft's 3D Viewer™.

Comparison of 2D-3D conversion results

RMS data on superimposition of 3D structures generated by ChemoSoft's 2D-3D Converter™ and original X-ray 3D ligand structures with bonding data were obtained by Strfit. The results were compared to those provided in the original paper [2] for 2D-3D conversion tools Corina™ and Concord™.

Creative pharmacophore search

The structures demonstrated the highest RMS values (from 0.89 till 1.70) for 2D-3D conversion were selected for further investigation by creative pharmacophore search (structures 6, 20, 26-32). At the first step a pharmacophore query was created from each ligand X-ray structure with the added bonding information (see above) by using ChemoSoft's Pharmacophore Query Editor™ (E3D) which is a part of ChemoSoft's 3D Tools™. First we tried to employ all the automatically detected pharmacophore centers and all distances between them. After that we tried to perform a pharmacophore search with the obtained query and the initial X-ray structure with bonding scheme as a file to search by using ChemoSoft's 3D Pharmacophore Search Engine™ which is a part of ChemoSoft's 3D Tools™ and consists of a command line utility Xpps and a GUI to it, called Searcher. The number of output conformers during creative search was set to 1. If the software had failed to find any hit (in case of too many pharmacophore centers) we sequentially removed pharmacophore centers and set all distances between the remaining centers. Then we tried to search for a pharmacophore hit with this new query. From hits obtained with different queries we selected one with the lowest RMS. Typical windows of ChemoSoft's Pharmacophore Query Editor™ and ChemoSoft's 3D Pharmacophore Search Engine™ are shown in Figs 3 and 4 correspondingly.

*Corresponding author. E-mail: an@chemdiv.com

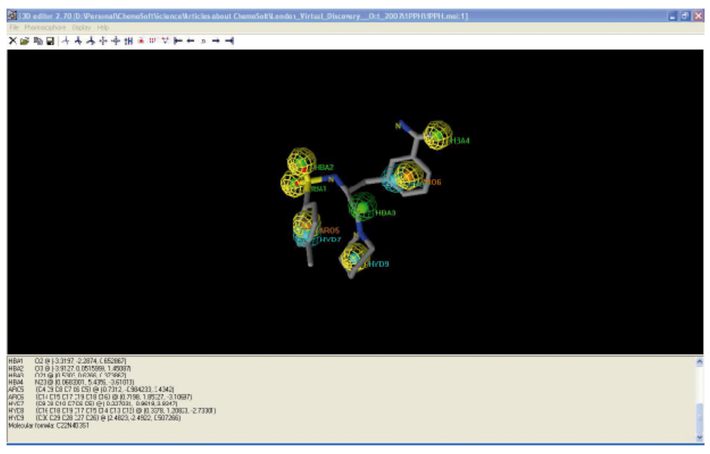


Figure 3. The window of ChemoSoft's Pharmacophore Query Editor™

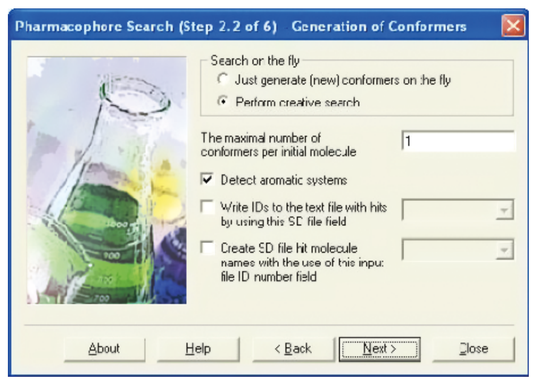


Figure 4. The window of ChemoSoft's 3D Pharmacophore Search Engine™

Comparison of creative pharmacophore search results

RMS data on superimposition of 3D hit structures creatively generated by ChemoSoft's 3D Pharmacophore Search Engine™ and original X-ray 3D ligand structures with bonding data were obtained by Strfit. The results were compared to those created by ChemoSoft's 2D-3D Converter™ and those provided in the original paper [2] for conformational search software tools Catalyst™, Confort™, and Omega™.

The chemical structures studied in this work

The molecular structures investigated in the present work were taken from the publication[2] and are shown in Table 1.

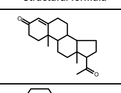
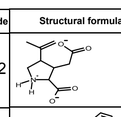
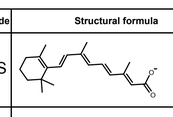
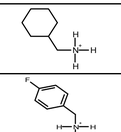
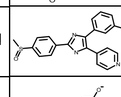
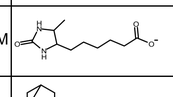
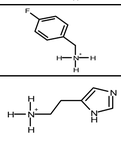
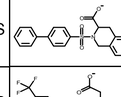
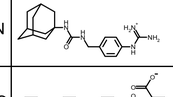
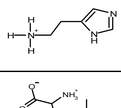
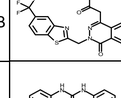
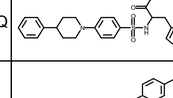
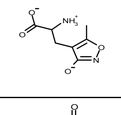
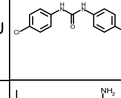
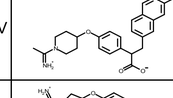
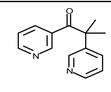
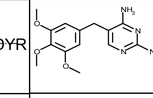
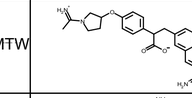
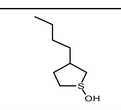
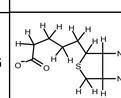
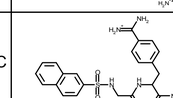
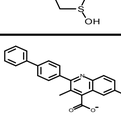
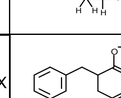
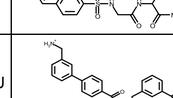
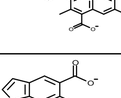
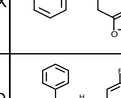
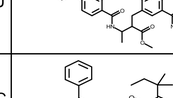
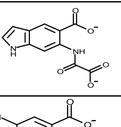
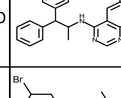
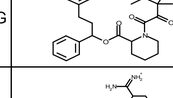
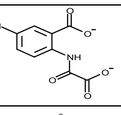
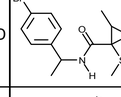
Structure number	PDB code	Structural formula	Structure number	PDB code	Structural formula	Structure number	PDB code	Structural formula
1	1A28		12	1GR2		23	1CBS	
2	1TNG		13	1IAN		24	1DAM	
3	1TNH		14	1BZS		25	1EJN	
4	1QFT		15	1FRB		26	1CAQ	
5	1FTM		16	1BJU		27	1MTV	
6	1PHG		17	1DYR		28	1MTW	
7	3BTO		18	2IZG		29	1PPC	
8	1D3G		19	1CBX		30	1F0U	
9	1C83		20	5STD		31	1FKG	
10	1ECV		21	6STD		32	1PPH	
11	1FCZ		22	7STD		*Structures 29 and 32 changed places in error in the original paper[2]		

Table 1. Structures studied in this work[2]*

Results and discussion

2D-3D Conversion

As highlighted in red in Table 2, ChemoSoft's 2D-3D Converter™ gave the best quality 3D structure conformation in 28 cases of 32 (87.5%) as compared to both Corina™ and Concord™.

Structure number	The number of rotatable bonds	RMS for 2D-3D conversion		
		ChemoSoft (1 conf.)	Corina[2]	Concord[2]
1	1	0,46	0,42	0,34
2	1	0,02	0,12	0,12
3	1	0,13	0,22	0,22
4	2	0,26	0,12	0,12
5	3	0,43	1,18	0,95
6	3	0,94	1,48	1,67
7	3	0,49	0,91	0,95
8	3	0,29	0,57	0,51
9	4	0,20	0,39	0,44
10	4	0,19	0,38	0,49
11	4	0,44	0,91	0,69
12	4	0,58	0,63	1,34
13	4	0,27	0,34	0,47
14	4	0,76	1,64	1,72
15	4	0,73	1,17	1,44
16	5	0,44	0,50	0,58
17	5	0,48	1,01	1,10
18	5	0,60	1,00	0,87
19	5	0,34	0,60	0,97
20	5	1,13	2,16	1,76
21	5	0,70	2,07	0,89
22	5	0,67	2,11	0,55
23	5	0,73	0,40	0,83
24	6	0,55	1,00	0,91
25	7	0,42	0,58	0,54
26	8	0,89	1,91	1,61
27	8	1,37	2,89	2,63
28	8	1,53	2,75	3,01
29	8	1,70	2,09	2,79
30	10	1,40	2,59	2,78
31	11	1,38	2,47	2,40
32	11	1,24	2,71	3,55

the lowest RMS (the best quality) of a 3D structure

Table 2. 2D-3D Conversion by ChemoSoft vs. competitors

This result clearly demonstrated that our low-energy conformation 2D-3D conversion approach seems a better alternative relative to traditionally employed rule-based techniques implemented in Corina™ and Concord™. Furthermore, as it can be seen from Table 3, the absolute quality of generated 3D structures generated by ChemoSoft's 2D-3D Converter™ are satisfactory in a significantly broader interval of molecular complexity (the number of rotatable bonds) as compared to the competitors. Also it is clear that the quality of conformers generated by ChemoSoft's 2D-3D Converter™ are considerably more consistently dependent on the number of rotatable bonds (molecular complexity) than those obtained by the other tools which can generate a bad conformation for almost any molecule, which makes it difficult to predict any relative outcome of 2D-3D conversion.

Structure number	The number of rotatable bonds	RMS for 2D-3D conversion		
		ChemoSoft (1 conf.)	Corina[2]	Concord[2]
1	1	0,46	0,42	0,34
2	1	0,02	0,12	0,12
3	1	0,13	0,22	0,22
4	2	0,26	0,12	0,12
5	3	0,43	1,18	0,95
6	3	0,94	1,48	1,67
7	3	0,49	0,91	0,95
8	3	0,29	0,57	0,51
9	4	0,20	0,39	0,44
10	4	0,19	0,38	0,49
11	4	0,44	0,91	0,69
12	4	0,58	0,63	1,34
13	4	0,27	0,34	0,47
14	4	0,76	1,64	1,72
15	4	0,73	1,17	1,44
16	5	0,44	0,50	0,58
17	5	0,48	1,01	1,10
18	5	0,60	1,00	0,87
19	5	0,34	0,60	0,97
20	5	1,13	2,16	1,76
21	5	0,70	2,07	0,89
22	5	0,67	2,11	0,55
23	5	0,73	0,40	0,83
24	6	0,55	1,00	0,91
25	7	0,42	0,58	0,54
26	8	0,89	1,91	1,61
27	8	1,37	2,89	2,63
28	8	1,53	2,75	3,01
29	8	1,70	2,09	2,79
30	10	1,40	2,59	2,78
31	11	1,38	2,47	2,40
32	11	1,24	2,71	3,55

Color coding:

RMS range	Color
≤0.50	
0.50-1.00	
1.00-1.50	
1.50-2.00	
>2.00	

Table 3. Quality of 2D-3D conversion

Moreover, as it follows from Table 4, the number of 3D structures obtained by ChemoSoft's 2D-3D Converter™ regularly decreases from the best to the worst quality range, which is not the case for the other two approaches.

RMS range	The number of investigated molecules per RMS range		
	ChemoSoft (1 conf.)	Corina[2]	Concord[2]
≤0.50	15	9	7
0.50-1.00	10	8	12
1.00-1.50	5	4	3
1.50-2.00	2	2	4
>2.00	0	9	6

Color coding:

The number of molecules in a column	Color
maximum	
...	
...	
...	
minimum	

Table 4. The number of hits per quality range by 2D-3D conversion

As clearly seen from Table 5, the relative number of good quality 3D structures is considerably higher for ChemoSoft™ as compared to Corina™ and Concord™. The opposite picture is for bad conformers.

RMS range	The number of investigated molecules per RMS range		
	ChemoSoft (1 conf.)	Corina[2]	Concord[2]
≥0.50	15	9	7
0.50-1.00	10	8	12
1.00-1.50	5	4	3
1.50-2.00	2	2	4
>2.00	0	9	6

Color coding:

The number of molecules in a row	Color
maximum	
...	
minimum	

Table 5. The number of hits by ChemoSoft vs. competitors

Creative pharmacophore search

As highlighted in red in Table 6, creative pharmacophore search by ChemoSoft's 3D Pharmacophore Search Engine™ gave the best quality 3D structure conformation in 9 "difficult" cases of 9 (100%) as compared to Catalyst™, Confort™, and Omega™.

Structure number	The number of rotatable bonds	RMS						
		ChemoSoft (2D-3D)	ChemoSoft (creative pharmacophore search)	Catalyst[2]		Confort[2]		Omega[2]
				Fast	Best	intermediate energy filter	final energy filter	Beam
6	3	0,94	0,05	0,66	0,55	1,57	1,72	0,15
20	5	1,13	0,25	0,79	0,77	1,11	2,53	0,87
26	8	0,89	0,78	1,01	1,26	0,81	2,57	1,99
27	8	1,37	0,41	0,86	1,09	1,23	2,13	1,29
28	8	1,53	0,52	0,98	0,84	1,57	2,08	1,52
29	8	1,70	0,51	0,81	0,96	2,16	1,70	0,63
30	10	1,40	0,78	0,94	1,20	1,99	1,37	2,01
31	11	1,38	0,91	1,17	1,31	1,34	2,62	1,12
32	11	1,24	0,69	1,48	1,44	1,45	0,92	0,87

- the lowest RMS (the best quality) of a conformation

Table 6. Conformations generated by creative pharmacophore search vs. other methods

As it follows from Table 7, the methodology used by creative pharmacophore search did adjust generated conformations to bioactive conformations employed to create pharmacophore queries as seen from the shift in the quality range by1-2 steps when moving from the conformations generated by ChemoSoft's 2D-3D Converter™ to those obtained during creative pharmacophore search. Moreover the conformers generated by creative pharmacophore search are often 1-4 RMS quality steps better than those obtained by using conformer generation tools Catalyst™, Confort™, and Omega™.

Structure number	The number of rotatable bonds	RMS						
		ChemoSoft (2D-3D)	ChemoSoft (creative pharmacophore search)	Catalyst[2]		Confort[2]		Omega[2]
				Fast	Best	intermediate energy filter	final energy filter	Beam
6	3	0,94	0,05	0,66	0,55	1,57	1,72	0,15
20	5	1,13	0,25	0,79	0,77	1,11	2,53	0,87
26	8	0,89	0,78	1,01	1,26	0,81	2,57	1,99
27	8	1,37	0,41	0,86	1,09	1,23	2,13	1,29
28	8	1,53	0,52	0,98	0,84	1,57	2,08	1,52
29	8	1,70	0,51	0,81	0,96	2,16	1,70	0,63
30	10	1,40	0,78	0,94	1,20	1,99	1,37	2,01
31	11	1,38	0,91	1,17	1,31	1,34	2,62	1,12
32	11	1,24	0,69	1,48	1,44	1,45	0,92	0,87

Color coding:

RMS range	Color
≤0.50	
0.50-1.00	
1.00-1.50	
1.50-2.00	
>2.00	

Table 7. The quality of conformations generated by creative pharmacophore search vs. other methods

Conclusions

1. The low-energy conformation approach implemented in ChemoSoft's 2D-3D Converter™ is a considerably better first approximation of actual bioactive conformations as compared to rule-based 2D-3D molecule conversion tools Corina™ and Concord™, at least for the studied dataset of chemical structures.
2. The good conformation quality range vs. complexity (the number of rotational bonds) is consistently broader for ChemoSoft's 2D-3D Converter™. It provides conformations consistently dependent on the number of rotatable bonds, thus, enabling qualitative predictions of 2D-3D conversion outcome.
3.ChemoSoft's 2D-3D Converter™ gives the highest number of good quality bioactive conformations as compared to Corina™ and Concord™. The number molecules per quality range regularly increases with the increase in quality for ChemoSoft's 2D-3D Converter™, which support the idea that low-energy conformation approach is bioactive conformation-oriented.
4. The creative pharmacophore search methodology implemented in ChemoSoft's 3D Pharmacophore Search Engine™ does adjust generated conformations to bioactive conformations from which corresponding pharmacophore queries are created.
5. The conformations generated with the use of the creative pharmacophore search targeted approach are of from considerably to dramatically better quality than those obtained by the other techniques (Catalyst™, Confort™, and Omega™) including conformational diversity-oriented approach (Catalyst™), at least for the investigated complex chemical structures.

1. www.chemosoft.com
2. Bostrom J. // J. Comp.-Aided. Mol. Design, (2001), V.15, P.1137-1152
3. Perola E., Charifson P. S. // J. Med. Chem. (2004), V.47, P.2499-2510