First Principles Calculation of Electron Ionization Mass Spectra for Selected Organic Drug Molecules

Christoph Alexander Bauer and Stefan Grimme*

Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie der Rheinischen Friedrich-Wilhelms-Universität Bonn, Beringstr. 4, D-53115, Bonn, Germany

E-mail: grimme@thch.uni-bonn.de

Phone: +49 (0)228/73-2351. Fax: +49 (0)228/73-9064

Contents

*To whom correspondence should be addressed

1 IUPAC names of compounds 1-5

1, Valsartan, 1*H*-tetrazol-tautomer: (*S*)-3-methyl-2-(N-[2'-(1*H*-1,2,3,4-tetrazol-5-yl)biphenyl-4-yl]methylpentanamido)butanoic acid.

2, Erythromycin: (3*R*,4*S*,5*S*,6*R*,7*R*,9*R*,11*R*,12*R*,13*S*,14*R*)-6- [(2*S*,3*R*,4*S*,6*R*)-4-(dimethylamino)-3-hydroxy-6-methyloxan-2-yl]oxy- 14-ethyl-7,12,13-trihydroxy-4-[(2*R*,4*R*,5*S*,6*S*)- 5-hydroxy-4-methoxy-4,6-dimethyloxan-2-yl]oxy- 3,5,7,9,11,13-hexamethyl-1-oxacyclotetradecane-2,10-dione.

3, Taxol: (2α,4α,5β,7β,10β,13α)-4,10-bis(acetyloxy)-13-[(2*R*,3*S*)- 3-(benzoylamino)-2-hydroxy-3-phenylpropanoyl]oxy- 1,7-dihydroxy-9-oxo-5,20-epoxytax-11-en-2-yl benzoate.

4, Lovastatin: (1*S*,3*R*,7*S*,8*S*,8a*R*)-8-2-[(2*R*,4*R*)-4-hydroxy-6-oxotetrahydro-2*H*-pyran-2-yl]ethyl-3,7-dimethyl-1,2,3,7,8,8a-hexahydronaphthalen-1-yl (2*S*)2-methylbutanoate.

5, Simvastatin: (1*S*,3*R*,7*S*,8*S*,8a*R*)-8-2-[(2*R*,4*R*)-4-hydroxy-6-oxotetrahydro-2*H*-pyran-2-yl]ethyl-3,7-dimethyl-1,2,3,7,8,8a-hexahydronaphthalen-1-yl 2,2-dimethylbutanoate.

2 XYZ Files of compounds 1-5 (optimized geometries) and Fragments 1a-1c (mean geometries)

The geometries below were optimized at the TPSS-D3/def2-TZVP level of theory (see article). For **3** there was one negative eigenvalue of the hessian matrix corresponding to an imaginary frequency at ca. -10/cm. This is regarded negligible for our purpose.

1, Valsartan, 1*H*-tetrazol-tautomer

```
61
Energy = -1432.318818848
```

O	0.5860684	3.9182096	-2.3901809
C	-0.5553080	3.5877570	-2.6360450
O	-1.4459527	4.4453467	-3.2081556
C	-1.1584245	2.2085010	-2.4090437
H	-1.2718155	1.7632319	-3.4070787
N	-0.1766088	1.3389239	-1.7557392
C	0.1903726	0.1982948	-2.4511047
O	-0.3574381	-0.0978048	-3.5136216
C	1.2555157	-0.7086297	-1.8489343
C	2.2241091	-1.2617646	-2.9080909
C	3.3431513	-0.2826573	-3.2924469
C	2.8567957	1.0326129	-3.9123407
C	0.4661093	1.7537297	-0.5073508
C	0.2134041	0.8409202	0.6777682
C	-0.8554378	-0.0577882	0.7198735
C	-1.0308754	-0.9066119	1.8100987
C	-0.1417282	-0.8761003	2.8927733
C	-0.2616087	-1.8689282	3.9946476
C	-0.7483747	-1.4710190	5.2443735
C	-0.9102592	-2.3894578	6.2793831
C	-0.5846128	-3.7311691	6.0740222
C	-0.0842364	-4.1430674	4.8450474
C	0.0912696	-3.2229001	3.7975691
C	0.6681362	-3.7315379	2.5502128
N	1.2592005	-3.0266582	1.5542982
N	1.6997871	-3.8764038	0.5988003
N	1.3827871	-5.0685319	1.0121266
N	0.7476936	-5.0190650	2.2130229
C	0.9064016	0.0553517	2.8711578
C	1.0833527	0.8956230	1.7729229
C	-2.5691941	2.1677144	-1.7454103
C	-2.6867416	3.0090923	-0.4680374
C	-3.6909616	2.5057176	-2.7382451
H	-0.9589914	5.2811656	-3.3481593
H	0.7115567	-1.5439641	-1.3903113
H	1.8146225	-0.2193807	-1.0461476
H	1.6414960	-1.5396015	-3.7928924
H	2.6723896	-2.1784393	-2.5093147
H	4.0124647	-0.7887858	-3.9996579
H	3.9489992	-0.0638227	-2.4011769
H	2.2782418	1.6334179	-3.2017904
H	2.2176434	0.8434672	-4.7813059
H	3.7047407	1.6454892	-4.2367182
H	0.0995403	2.7555584	-0.2762774
H	1.5453557	1.8677514	-0.6575614

```
H    -1.5351476   -0.1187090   -0.1238982
H    -1.8457020   -1.6251301    1.8119293
H    -1.0173779   -0.4288168    5.3911553
H    -1.2945793   -2.0593699    7.2400467
H    -0.7123020   -4.4543139    6.8738128
H     0.1903331   -5.1780971    4.6709458
H     1.3783332   -2.0247978    1.4538608
H     1.6064795    0.0897275    3.7014640
H     1.9223502    1.5879542    1.7571197
H    -2.6922487    1.1113449   -1.4765594
H    -1.9973891    2.6769033    0.3137665
H    -3.7010217    2.9306409   -0.0627956
H    -2.4948981    4.0677785   -0.6768763
H    -3.6013709    1.9035867   -3.6485366
H    -3.6765424    3.5607024   -3.0202975
H    -4.6615522    2.2868090   -2.2800959
```

2, Erythromycin

```
118
Energy = -2484.369464843
C    -3.4800253   -0.3003184    4.8314421
C    -2.8282087   -1.5431512    4.2176657
C    -3.0385018   -1.6104939    2.7063514
H    -4.1055328   -1.4802330    2.4791660
O    -2.3693927   -0.4315361    2.1489531
C    -2.9753114    0.2291172    1.1385975
O    -3.9425164   -0.2170888    0.5268774
C    -2.4466866    1.6516977    1.0286751
H    -1.9042441    1.8207994    1.9640963
C    -3.6615967    2.6019468    1.0293308
C    -1.3712746    2.0336359   -0.0431107
H    -1.3837679    3.1321806    0.0337183
O    -0.0838139    1.5937052    0.4268974
C     0.8769356    2.6616958    0.5803411
H     0.8656367    3.2766908   -0.3270800
O     2.1539013    2.0872894    0.6216159
C     2.3311997    1.0988439    1.6804181
H     1.8704068    0.1677440    1.3314403
C     3.8280803    0.8826633    1.8375675
C     1.5979886    1.5043031    2.9855034
H     2.1980366    1.1976461    3.8492262
C     1.3959229    3.0357077    3.0602087
C     0.5837802    3.5103395    1.8369181
C     2.7440246    3.7619172    3.1053380
O     0.7833941    3.4623913    4.2972492
```

C	-0.5029977	2.9674231	4.6723463
O	0.3658387	0.7844126	3.1031272
C	-1.5432294	1.7545452	-1.5703356
H	-1.0785822	2.6337168	-2.0346866
C	-3.0123638	1.8171467	-2.0187808
C	-0.7612933	0.5942734	-2.2542055
H	-1.0769397	0.6212617	-3.3062257
O	0.6627462	0.8437431	-2.2462140
C	1.1109131	1.9120155	-3.0580417
H	0.6148579	2.8518939	-2.7852492
O	0.7864956	1.7719632	-4.4369394
C	1.5434555	0.7892971	-5.2090336
H	2.0770519	1.3720883	-5.9750640
C	0.5436670	-0.1330621	-5.8930146
C	2.5629369	0.0382588	-4.3592904
C	3.4615328	1.0168303	-3.6029963
H	4.0228694	1.5991749	-4.3475728
C	2.6131313	2.0644499	-2.8152428
H	2.7839584	1.9696912	-1.7407376
O	3.0016845	3.4128757	-3.1411499
N	4.4777166	0.3141946	-2.8069402
C	5.4370942	1.2475519	-2.2149730
C	3.9229278	-0.5902316	-1.7975330
C	-1.0205525	-0.8440211	-1.7382048
C	0.1735330	-1.7820880	-2.0084610
C	0.0445132	-3.1656778	-1.3472324
H	-0.9378920	-3.6097924	-1.5330743
C	0.2860233	-3.1135757	0.1582263
O	1.1373109	-2.3602130	0.6254484
C	-0.3962241	-4.1116972	1.1121714
H	0.4647638	-4.5531625	1.6322362
C	-1.1591972	-3.4101054	2.2921172
H	-1.2777207	-4.1935551	3.0538193
C	-2.6258239	-2.9243016	1.9925545
C	-3.6291289	-4.0135356	2.4059382
O	-2.7853530	-2.7083929	0.5637494
O	-0.3547173	-2.4110850	2.8869198
C	-1.1822835	-5.2511056	0.4545169
C	1.1244223	-4.1080222	-1.9362674
C	-2.2934258	-1.3925503	-2.4100571
O	-1.1957074	-0.7552372	-0.3249335
H	-3.3515741	-0.2978143	5.9182412
H	-4.5552485	-0.2636056	4.6173624
H	-3.0192361	0.6101199	4.4380410
H	-3.2668501	-2.4453459	4.6600325

H	-1.7564254	-1.5559724	4.4209544
H	-4.3512836	2.3828971	0.2144470
H	-3.3246799	3.6402639	0.9449778
H	-4.2112417	2.4993803	1.9704630
H	4.3291439	1.7643815	2.2430780
H	4.2748384	0.6432950	0.8696438
H	3.9986210	0.0391300	2.5141901
H	-0.4872851	3.4662398	2.0531377
H	0.8286752	4.5588038	1.6376650
H	3.3743230	3.3530767	3.9011988
H	2.5557221	4.8168396	3.3254538
H	3.2632224	3.6854665	2.1482330
H	-0.4569995	1.9398502	5.0418947
H	-1.2217595	3.0025508	3.8419927
H	-0.8461446	3.6411248	5.4631597
H	-0.0261960	0.7401128	2.2069136
H	-3.4268212	2.8125017	-1.8280825
H	-3.6353840	1.0871988	-1.4990417
H	-3.0809678	1.6351932	-3.0964922
H	1.0583116	-0.8009872	-6.5929957
H	-0.1959866	0.4512962	-6.4482839
H	0.0192048	-0.7442126	-5.1511470
H	2.0317081	-0.6155399	-3.6650060
H	3.1804131	-0.5885860	-5.0121689
H	2.7284082	3.5716985	-4.0607269
H	5.0002257	1.9079105	-1.4446603
H	5.8552359	1.8813032	-3.0032226
H	6.2484300	0.6748145	-1.7538230
H	3.2237909	-0.1023044	-1.0978535
H	4.7517007	-1.0108448	-1.2201174
H	3.3980687	-1.4175684	-2.2811391
H	0.2958567	-1.9218699	-3.0898143
H	1.0772715	-1.3084667	-1.6251669
H	-3.5502083	-4.2469350	3.4719447
H	-3.4418953	-4.9271027	1.8387904
H	-4.6513353	-3.6828345	2.1943048
H	-3.4976211	-2.0318963	0.4511592
H	0.2922637	-2.0881673	2.2224006
H	-2.0663951	-4.8897905	-0.0728855
H	-1.5047073	-5.9660757	1.2192207
H	-0.5532030	-5.7991180	-0.2527306
H	2.1159108	-3.6692037	-1.7895119
H	0.9536107	-4.2386596	-3.0092144
H	1.1144871	-5.0958402	-1.4653429
H	-2.1648016	-1.4355487	-3.4969516

```
H    -3.1590123   -0.7691176   -2.1839990
H    -2.5102243   -2.4031674   -2.0515579
H    -1.6804036   -1.5617631   -0.0250857
```

3, Taxol

```
113
Energy = -2931.314997945
C    -0.4225808   -2.9908059   -1.3526640
C    -1.3102084   -2.0773828   -0.5593038
C    -2.4315961   -1.4657497   -0.9981728
C    -2.8171082   -1.5128167   -2.4413512
H    -2.0401911   -2.0084352   -3.0326659
C    -3.0308586   -0.1638558   -3.1684469
O    -3.9634070   -0.0964918   -3.9572107
C    -2.0032895    0.9913640   -3.0219198
C    -2.2111527    1.9061759   -4.2767153
H    -2.0359282    1.2646739   -5.1525951
C    -1.2467635    3.0793983   -4.2793488
C    -1.6181403    3.9956694   -3.1293123
H    -2.1622338    4.8789661   -3.4641460
C    -2.2990587    3.3859338   -1.8794372
C    -2.2849497    1.8141529   -1.6919701
H    -3.3417991    1.6604103   -1.4984918
C    -1.6780447    1.4856586   -0.2940325
H    -2.0103159    2.3508439    0.2804997
C    -2.2804117    0.3429835    0.6446551
C    -3.2871150   -0.6691782   -0.0137519
C    -4.5715772   -0.0307602   -0.5950372
C    -3.8806235   -1.5945985    1.0937051
C    -1.2256071   -0.4742719    1.4303003
C    -0.9381564   -1.8886501    0.8933953
H    -1.4467267   -2.6509290    1.4897483
O     0.4879965   -2.1474037    1.1077951
C     0.8258576   -3.1828742    1.8823232
O     0.0647077   -4.0146044    2.3516495
C     2.3238025   -3.2225655    2.1328708
H     2.8605554   -2.7729418    1.2971847
C     2.6649929   -2.4437322    3.4466309
H     2.2918619   -3.0867450    4.2504458
C     1.9890608   -1.0928786    3.6066148
C     2.4218598    0.0524560    2.9281386
C     1.7587747    1.2659272    3.1030391
C     0.6574614    1.3541860    3.9564608
C     0.2217661    0.2181249    4.6375980
C     0.8842169   -0.9969804    4.4612683
```

```
N     4.1223598   -2.4148884    3.6277011
C     5.0547171   -2.0598881    2.6969760
O     4.7738271   -1.5419319    1.6110378
C     6.4776395   -2.3232152    3.0898547
C     7.4650539   -1.4904039    2.5503506
C     8.8047277   -1.6915326    2.8690166
C     9.1732628   -2.7403105    3.7147456
C     8.1968849   -3.5885224    4.2377115
C     6.8533597   -3.3793071    3.9301073
O     2.7369979   -4.5768488    2.2685754
O    -2.9942387    1.1718248    1.6089159
O    -0.2304214    1.7399276   -0.2223430
C     0.8173912    0.9251925   -0.4975783
O     0.7449186   -0.2671835   -0.7127909
C     2.1097383    1.6601476   -0.4672764
C     2.2010031    3.0534338   -0.5757638
C     3.4504591    3.6684272   -0.5422911
C     4.6072211    2.9015119   -0.3866581
C     4.5179009    1.5110780   -0.2849733
C     3.2745582    0.8905253   -0.3386711
C    -1.3025569    4.2730396   -1.1020333
O    -0.4812238    4.4072760   -2.2964652
O    -3.7117735    3.6866442   -1.6340720
C    -4.2505122    4.9249860   -1.7960292
O    -3.6379384    5.9334938   -2.0782475
C    -5.7443907    4.8478069   -1.5865868
O    -3.5162271    2.4865857   -4.3527001
C    -0.6072149    0.3608219   -3.1722702
O    -4.0337951   -2.2912802   -2.5903075
C    -4.1033569   -3.0547242   -3.7220030
O    -3.1826671   -3.1911843   -4.4990156
C    -5.4664100   -3.6814461   -3.8494390
H     0.5637241   -2.5383880   -1.4904575
H    -0.8405399   -3.2481650   -2.3271782
H    -0.2758061   -3.9251088   -0.7966609
H    -1.3313238    3.6042956   -5.2349529
H    -0.2131413    2.7462290   -4.1660081
H    -5.3613787   -0.7856960   -0.6228719
H    -4.5041050    0.3621933   -1.6055477
H    -4.8890767    0.7805324    0.0662334
H    -4.3978536   -2.4245873    0.6058272
H    -4.6282998   -1.0420669    1.6761231
H    -3.1557841   -2.0144792    1.7898844
H    -1.5918213   -0.5630411    2.4569696
H    -0.2867226    0.0689416    1.5011464
```

```
H     3.2724646   -0.0164235    2.2606272
H     2.1011267    2.1427052    2.5613444
H     0.1404703    2.3005163    4.0859091
H    -0.6304571    0.2765601    5.3093386
H     0.5414353   -1.8813443    4.9945512
H     4.4706491   -2.7838906    4.5012518
H     7.1581700   -0.6909462    1.8835240
H     9.5643736   -1.0347764    2.4549135
H    10.2196184   -2.9020401    3.9574854
H     8.4818414   -4.4198360    4.8760222
H     6.1053590   -4.0714614    4.3088473
H     1.9497336   -5.0545563    2.5960334
H    -3.3770238    0.5746871    2.2731688
H     1.3060733    3.6447313   -0.7198691
H     3.5205293    4.7477618   -0.6408379
H     5.5783661    3.3877117   -0.3509544
H     5.4101733    0.9057185   -0.1590765
H     3.1898876   -0.1869664   -0.2575411
H    -0.7421379    3.8498580   -0.2693300
H    -1.7561858    5.2295188   -0.8274765
H    -6.1479123    5.8548521   -1.4808694
H    -5.9828831    4.2408909   -0.7094124
H    -6.1953247    4.3646557   -2.4602288
H    -4.1176117    1.7173750   -4.4009001
H    -0.4439412   -0.4404645   -2.4578052
H     0.1789376    1.1016759   -3.0195813
H    -0.5021133   -0.0390986   -4.1873440
H    -5.4226963   -4.5101679   -4.5561263
H    -6.1597987   -2.9216975   -4.2263510
H    -5.8333587   -4.0185374   -2.8770473
```

4, Lovastatin

```
65
Energy = -1313.155067414
C    -1.0713715   -0.6945412    3.2175892
C     4.7183396   -2.8480415   -2.2345382
C     1.6460810   -2.1766252    3.6951827
C     1.3430963    2.4653235    0.1926034
C     0.3184623   -0.0686157    3.0724197
C     2.2993697    2.7416549   -2.0898992
C     3.1709699    1.7789900   -2.4282920
C    -1.9225971    1.6958411   -0.9738088
C    -0.8003841    0.6525726   -1.0083244
C     3.9436636   -0.4750741   -1.9023064
C     2.3463757   -2.3688286   -1.4846538
```

```
C    -3.7958095    0.0243189   -0.4617325
C    -5.9511767    1.0240322    0.2520028
C     3.8070731   -1.8976473   -1.4396645
C     1.4497697   -1.0549796    2.6695902
C     1.0455060    2.4267012   -1.3197672
C     2.9226263    0.3787114   -2.1101661
C    -5.0274532   -0.1795822    0.4102287
C    -3.0601036    1.2889412   -0.0445979
C     0.4343196    1.0912786   -1.8145554
C     1.3623498   -1.2807983   -1.0576787
C    -5.2637533    2.3796586    0.3359855
C     1.4704224   -0.0754960   -1.9947219
C     1.1484367   -1.6328363    1.2990561
O    -4.5497554   -0.3035116    1.7631471
O    -5.8708313    3.3948597    0.5934132
O     0.4991349   -2.6396049    1.0872420
O    -3.9459498    2.4612135    0.0190840
O     1.6623075   -0.8498844    0.3111774
H    -1.8094844    0.0742513    3.4665967
H    -1.0921395   -1.4470755    4.0119934
H    -1.3952814   -1.1811581    2.2931164
H     4.6377333   -3.8722170   -1.8530836
H     5.7667764   -2.5390475   -2.1621437
H     4.4379872   -2.8530500   -3.2939779
H     1.7886034   -1.7503673    4.6938594
H     2.5255262   -2.7803600    3.4499914
H     0.7795789   -2.8414827    3.7143903
H     1.8177094    3.4147349    0.4610539
H     2.0161595    1.6505595    0.4711271
H     0.4222228    2.3754287    0.7793099
H     0.6182363    0.3932842    4.0216405
H     0.2864830    0.7375658    2.3313501
H     2.5327438    3.7866656   -2.2835822
H     4.1261440    2.0316225   -2.8853864
H    -2.3212040    1.8629498   -1.9830981
H    -1.5431384    2.6586220   -0.6168322
H    -0.5038036    0.4077646    0.0177447
H    -1.1906884   -0.2735201   -1.4469928
H     4.9620064   -0.0978319   -2.0085256
H     2.0878640   -2.6553410   -2.5129321
H     2.2068640   -3.2549539   -0.8583073
H    -4.1000565    0.0881282   -1.5139846
H    -3.1201071   -0.8292928   -0.3524046
H    -6.7594415    1.0217092    0.9900559
H    -6.4311273    0.9860503   -0.7356679
```

```
H     4.1439043   -1.9274396   -0.3901873
H     2.3698278   -0.4671835    2.5788417
H     0.3089965    3.2122483   -1.5304903
H    -5.5552352   -1.0989440    0.1179784
H    -2.6782746    1.1558289    0.9732734
H     0.0827302    1.3180046   -2.8297899
H     0.3507547   -1.6906683   -1.0476013
H     1.2201032   -0.5110590   -2.9735766
H    -5.3180924   -0.3624564    2.3522261
```

5, Simvastatin

```
68
Energy = -1352.501082429
C    -2.1418192    0.3427751   -1.6308656
C     0.5699329   -5.1374901    2.7487755
C    -1.2407826    2.5337091    2.7961130
C    -1.7514921   -3.5508669   -1.7803704
C    -1.0403112   -1.8527576   -3.5026095
C    -2.4279693   -1.1527870   -1.4787564
C    -0.6161777    0.6448596    4.2162892
C    -0.7015694   -0.6941073    4.2515023
C     0.6564199    2.6167746    0.2126150
C     1.0067808    1.2128035    0.7200837
C    -0.4110407   -2.8614964    3.1568941
C     1.2989678   -3.1400544    1.3689974
C     1.3022763    4.4456277   -1.4390819
C     1.0162787    3.8537592   -3.8257691
C     0.1388399   -3.8018917    2.1231940
C    -0.1649516    1.4466133    3.0215056
C    -0.1914827   -1.5322556    3.1822751
C     1.8511678    4.6750820   -2.8479472
C     1.4341459    2.9828329   -1.0487234
C     0.0722378    0.5684620    1.7643538
C     0.9655433   -1.7129617    0.9489544
C     0.7568021    2.4160864   -3.3972430
C     0.6605360   -0.8176787    2.1531484
C    -0.0042513   -1.9075515   -1.2362056
C    -1.3123598   -2.0839522   -2.0086452
O     3.2527367    4.3626014   -2.9327833
O     0.3533925    1.5750695   -4.1689618
O     1.1081686   -1.9655066   -1.7203532
O     0.9524978    2.0671771   -2.0971989
O    -0.2269602   -1.7390968    0.0992589
H    -2.9470606    0.9334736   -1.1799358
H    -2.0521537    0.6344634   -2.6808964
```

```
H    -1.2048942    0.6143903   -1.1366562
H     0.9427316   -5.8234926    1.9801650
H    -0.2705412   -5.6213864    3.2584408
H     1.3677259   -4.9802953    3.4837968
H    -2.1861592    2.0642248    2.5024809
H    -0.9630184    3.2549328    2.0266629
H    -1.4155007    3.0870515    3.7255172
H    -1.9706725   -3.7338540   -0.7236444
H    -0.9726819   -4.2479910   -2.1063011
H    -2.6577252   -3.7509106   -2.3624769
H    -0.6346858   -0.8560221   -3.6910375
H    -1.9759139   -1.9673095   -4.0620409
H    -0.3174953   -2.5824496   -3.8765885
H    -3.3462805   -1.4098600   -2.0237075
H    -2.6070899   -1.3897559   -0.4247151
H    -0.9699284    1.2258305    5.0673467
H    -1.1355890   -1.1907203    5.1179329
H     0.8775970    3.3755265    0.9725750
H    -0.4109887    2.6918242   -0.0256316
H     1.0427189    0.5576424   -0.1549335
H     2.0263569    1.2376246    1.1337205
H    -1.0280387   -3.2989476    3.9435084
H     2.1804722   -3.1010508    2.0237112
H     1.5740097   -3.7244254    0.4846294
H     0.2487788    4.7532355   -1.4015194
H     1.8575475    5.0477804   -0.7128212
H     1.4731364    3.8235066   -4.8185729
H     0.0276152    4.3159154   -3.9448108
H    -0.6606588   -4.0194368    1.3972531
H     0.7817099    1.9545954    3.2744735
H     1.7964695    5.7355989   -3.1101168
H     2.4952026    2.7339491   -0.9164971
H    -0.9041543    0.3889847    1.2964334
H     1.7809194   -1.3068174    0.3486819
H     1.6355226   -0.6169893    2.6291067
H     3.3571591    3.3973850   -2.9619790
```

The next three geometries provided are heavily distorted mean fragment geometries as produced by QCEIMS. **1a-c** are radical cations.

1a, $C_{14}H_{10}$

```
        24

C              2.01865329        5.82013539        3.25146337
C              0.78540618        6.50393559        3.11639035
C              0.47947089        6.75426286        1.68943183
```

C	-0.38544575	7.82292658	1.49685453
C	-0.52509596	8.54276439	2.67430600
C	-0.56523364	9.96636472	2.62596679
C	-0.04307674	10.85202052	1.59178175
C	-0.71343987	12.10583738	1.82675215
C	-1.00093923	12.48459788	3.16764046
C	-1.33807384	11.66389862	4.25085458
C	-1.01614598	10.33795860	3.95312412
C	-1.68901128	9.16564424	4.52692247
C	-0.83552240	8.02403699	4.06990788
C	-0.15492867	6.80850452	4.23081687
H	2.28733646	5.07693047	4.01143700
H	2.76229462	6.04055498	2.46441096
H	0.89178693	6.23038072	0.88477428
H	-0.66872611	8.25776378	0.54133118
H	0.92382583	10.65078249	1.12142114
H	-0.86796713	12.88255951	1.06526763
H	-1.28974089	13.53137944	3.28011967
H	-1.88644667	11.94605899	5.06827226
H	-2.70083825	9.00724223	4.76586174
H	-0.11200616	6.23239562	5.14418227

1b, $C_{14}H_{10}$

24

C	-8.60757956	19.84367775	-5.12416083
C	-8.47928060	18.84385609	-6.13404479
C	-9.64824089	18.92492690	-7.16166280
C	-9.87138822	18.02746478	-8.19996534
C	-8.90770297	16.97842361	-8.35947579
C	-8.69653387	16.20070657	-9.62547750
C	-7.79597115	16.25577654	-10.72246775
C	-7.52921949	15.08527227	-11.53604092
C	-7.63673814	13.67488678	-11.36307373
C	-8.84271782	13.05055036	-11.02324949
C	-9.82408721	15.20874866	-9.75071163
C	-9.53798522	14.16028643	-10.45375255
C	-7.97060066	16.74680721	-7.27772840
C	-7.64773917	17.61449766	-6.20806512
H	-9.35009821	20.49492944	-5.14911915
H	-7.95062548	19.93691631	-4.33043360
H	-10.33552586	19.75129568	-7.01642691
H	-10.72056256	18.08945573	-8.88599866
H	-7.22554767	17.17146346	-10.96706264
H	-6.75719956	15.29049390	-12.25986657

```
H            -6.69298381       13.09492377      -11.51538159
H            -9.25033373       12.07097330      -11.24263144
H            -7.63220548       15.70524865       -7.19661906
H            -6.80477247       17.48019740       -5.69891621
```

1c, $C_{14}H_{10}$

```
24

C            -9.81146003       10.36971879       -6.76624450
C           -10.90098711        9.43599084       -6.74311789
C           -12.14729944        9.58294303       -7.45387994
C           -12.88882419        8.36271252       -7.83959971
C           -12.91230351        7.16953755       -7.00132581
C           -13.13317000        4.68905955       -6.50671043
C           -13.66128037        3.48562777       -5.98536358
C           -14.47047476        2.54282335       -6.62805047
C           -15.25930783        3.16276417       -7.58924409
C           -15.65100769        4.58834357       -7.76086830
C           -13.71740589        5.92758177       -6.87856657
C           -15.09789550        5.77093712       -7.23722073
C           -11.56982898        6.97641199       -6.49344358
C           -10.74442295        8.05922709       -6.21467993
H            -9.92837318       11.22446583       -7.14371202
H            -8.81563944       10.04028849       -6.40849564
H           -12.71593838       10.51994653       -7.48353351
H           -13.66785960        8.47747521       -8.61254008
H           -13.13845189        3.20158130       -5.04673300
H           -14.56421049        1.49747417       -6.34334585
H           -15.56466991        2.51326207       -8.43140119
H           -16.20676266        4.74934980       -8.56649208
H           -11.24855317        5.98743846       -6.26200182
H            -9.80776891        7.81662504       -6.06126552
```

3 QCEIMS in more detail

For a comprehensive description of the inner workings of QCEIMS, the reader is referred to the SI of the original paper.[1] There, one of the authors (Grimme) has laid out the general QCEIMS architecture and discussed many technical details. What follows in this document is a more detailed overview of the procedure used for this study. The QCEIMS program was used in 3 steps:

1. **Initialization** - Production of a ground state trajectory to yield an ensemble from which to take snapshots (nuclear geometries and velocities) for the fragmentation runs. Optimized geometries were used as starting points. The velocity verlet algorithm[2] was used for propagation on the OM2-D3[3,4] potential energy surface (PES). For 1,000 trajectories, there were 25,000 steps of equilibration and 50,000 steps of production. The latter yielded the actual data from which the snapshots were obtained. With a timestep of 0.50 fs, this makes for 37.5 ps of ground state initialization. Geometry optimization at the given level of theory and initialization are actually the two most expensive steps in our procedure.

2. **Setup** - A random set of 1,000 equidistantly chosen nuclear geometry and velocity arrays was prepared and the ionization excess energy (IEE) was computed according to a Poisson distribution. Apart from $\varepsilon_{\mathrm{HOMO}}$ there was no information from the molecule itself. The distribution was calculated in such a way that the IEE/atom was equal to 0.6 eV.

3. **Production** - Each individual snapshot geometry was deprived of one electron, thereby rendering the systems radical cations. From this point, the propagation was started in the same way as for the closed-shell original system case, using either the OM2-D3 or the DFTB3-D3[5] PES. In the case of OM2-D3, unrestricted SCF calculations were carried out throughout every fragmentation run. The vibronic heating associated with internal conversion (IC) was simulated by scaling all nuclear velocities uniformly until the internal energy, predifined by the IEE value according to the Poisson distribution, was reached. This is a deviation from the protocol used in the original paper, where the velocities were scaled according to MO localizations (implications are discussed below). The heating time of the IC process was in the 1-2 ps time range. During the production runs, the 'electronic temperature' was set to a constant value of 21,000 K (OM2) or 5,000 K (DFTB3), yielding fractional occupation numbers (FON, 'Fermi Smearing').[6,7]

The FON approach had to be used in order to achieve SCF convergence.[8] Moreover, FON partially account for the mixing in of excited states and their multiconfigurational character. In the event of a fragmentation, a fragment assignment algorithm ordered the atoms according to their fragments. For each fragment, a mean geometry was computed from the trajectory. The ionization potentials (IPs) were then estimated at the OM2 level of theory for those mean fragment geometries. The charge was assigned statistically according to the Boltzmann distribution at the current (vibronic) temperature of the fragmentation event. The run was continued using the statistically most highly charged species, assigning the spin according to spin population analysis. This was done recursively up to a maximum of 7 times. The initial run maximum simulation time was 5 ps and decreased in each subsequent run. Given 4 or 5 runs during such a decomposition cascade, the typical total simulation time was between 5 and 10 ps.

At last, the fragments yielded by the production runs were counted with their statistical (Boltzmann-IP) weight and plotted vs. the experimental spectra (provided in either JCAMP-DX or peak table formats by the databases). The typical maximum count (100 % rel. int. signal in the computed spectra) was around 300. This is of course inferior to the experimental situation where there are millions of counts.

For OM2 calculations, the MNDO program[9] was called. DFTB+[5,10] was used for DFTB3 calculations.

The total computational times for spectra generation are found in Table 1. These are only rough estimates of the timescale for QCEIMS calculations. The number of available cores on our computer cluster ranged from 100 to 500 and Intel CPUs were used as well as AMD CPUs. Note that up to this point, the 37.5 ps of ground state initialization is by far the most expensive step in our procedure with a wall time of about 1 day for systems **1-5**. Table 1 also reveals that perhaps one of the strongest features of QCEIMS is the perfect

parallelization - each production run is conducted totally independently from all the others. Thus, predictions for **1-5**, which are at the limit of experimental EI-MS, are accessible within a few hours on a computer cluster and within a few days on a single workstation (assuming 8 modern CPU cores). It is also noteworthy that no part of the algorithm needs a lot of memory - provided semiempirical methods are used throughout. All in all, 8 to 16 GB RAM, which is standard nowadays, should be sufficient.

Table 1: Computational times (wall times) for QCEIMS production. Results marked (*) were produced on slow machines.

Compound	QC method	estimate single core 1000 runs	average prod. run	total time 1000 runs/ ca. 300 cores
1	OM2-D3	51 d	50 min	4 h 0 min
1	DFTB3-D3	124 d	40 min	2 h 0 min
2	OM2-D3	266 d	176 min	13 h 0 min
2	DFTB3-D3	138 d	66 min	3 h 0 min
3	OM2-D3	308 d	546 min*	70 h 0 min*
3	DFTB3-D3	132 d	77 min	6 h 0 min
4	OM2-D3	56 d	53 min	8 h 45 min
4	DFTB3-D3	93 d	36 min	1 h 45 min
5	OM2-D3	69 d	54 min	2 h 45 min
5	DFTB3-D3	105 d	34 min	2 h 45 min

4 Additional computational results

4.1 Velocity scaling effect

The effect of localized heating by MO populations during a QCEIMS production run is dramatic for larger systems. As seen from the spectral comparison displayed in Figure 1, a large number of artifacts results. This can be rationalized by inspecting the fragmentation runs. When heated locally, the Taxol frame breaks in certain positions very quickly, while other regions of the molecule stay 'cold'. If the charge is assigned to these 'cold parts', the runs are continued with these fragments, which then do not decompose further, leading to signals in the m/z 400-850 region. From the comparison with the experimental data one can see that this is wrong. While scaling the nuclear velocities uniformly leads to the reported computed spectrum of good quality, one cannot exclude other, non-beneficial effects of disregarding the localization of the ionization completely.

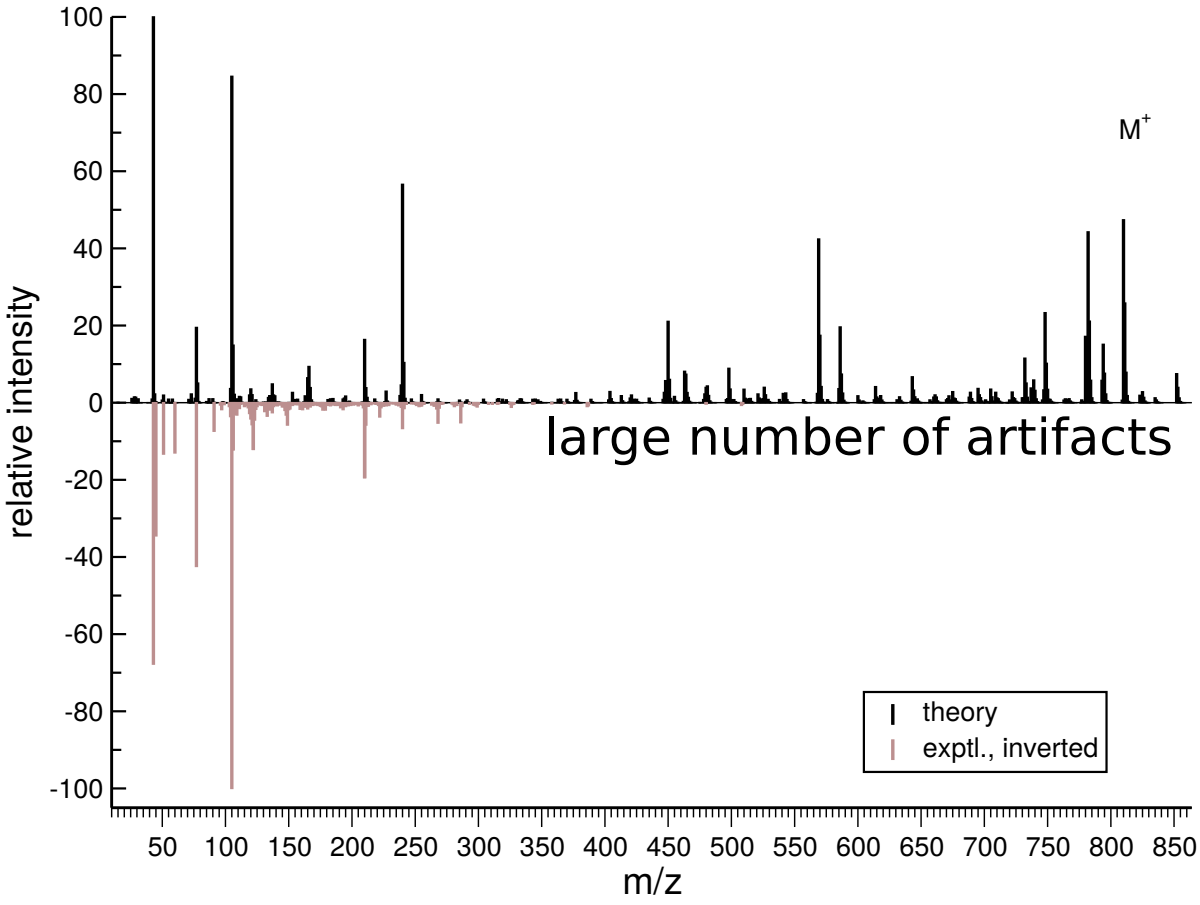

Figure 1: Computed (OM2-D3) vs. experimental spectrum of taxol (**3**) with local nuclear velocity scaling switched on. Note the artifacts in the heavier fragment region.

On the other hand, the computed spectra for other (smaller) compounds actually get better when the localization algorithm is switched on. However, in order to stay consistent, every reported spectrum (except where clearly noted) was produced with the localization routine deactivated.

4.2 IP estimate by OM2 - Comparison with DFT

Figure 2 shows the difference for two computed spectra of **5**. For the first spectrum, fragment IPs were computed at the PBE0/SVx level of theory and for the second spectrum, OM2 was

used for fragment IP calculation. The difference is marginal, with the maximum difference lower than 5 relative intensity units. One may thus safely assume that the semiempirical OM2 method is an adequate and efficient choice to gain a suitable IP estimate during a QCEIMS run.

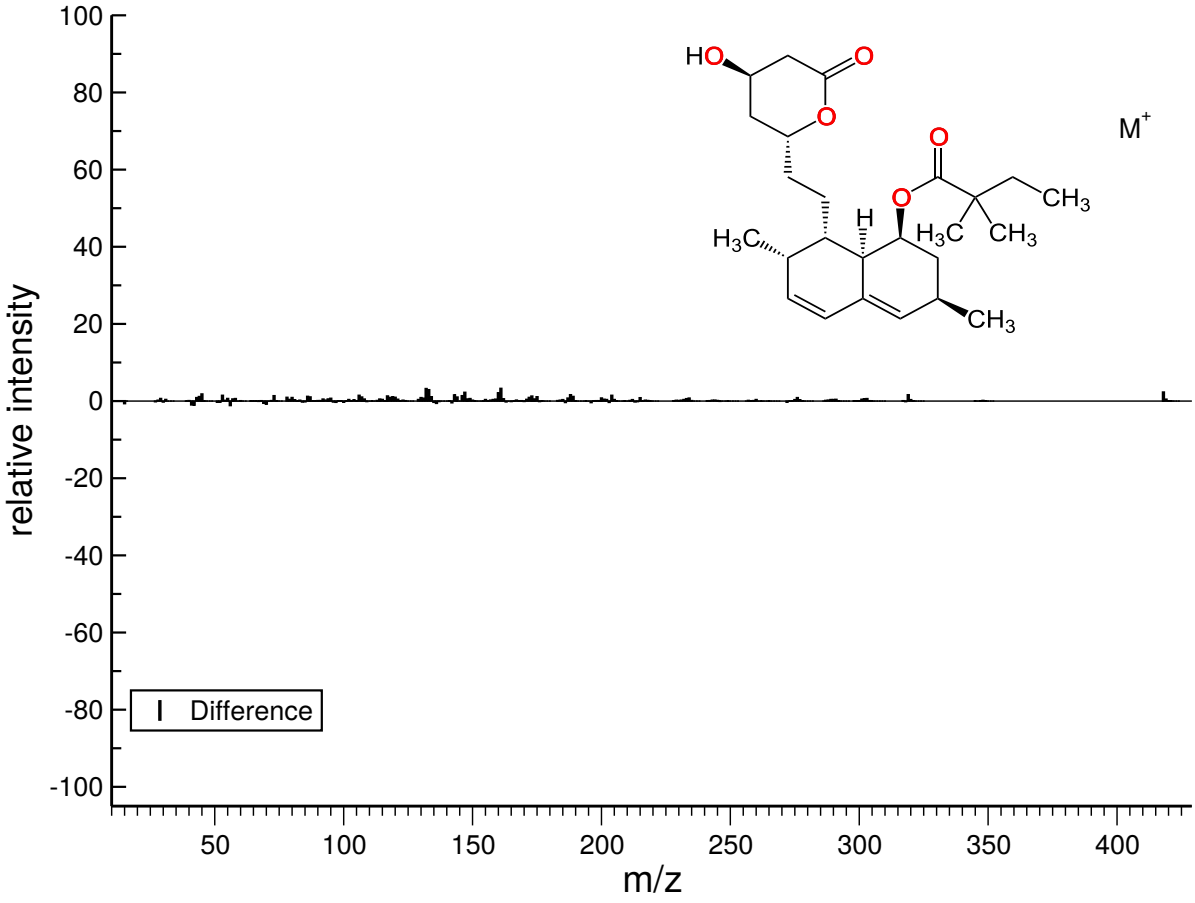

Figure 2: Difference between two computed spectra for Simvastatin: IP calculation with DFT (PBE0/SVx) - IP calculation with OM2.

4.3 DFTB3 spectra for 1-5

EI-MS spectra have been computed for compounds **1-5** using the DFTB3-D3 PES. As can be seen in Figures 3-7, they are of similar or slightly worse quality to the OM2-D3 results. This shows that QCEIMS is in principle workable for many QC methods, producing results according to the PES generated by the method. In our cases, both DFTB3-D3 and OM2-D3

work well. The reason why OM2-D3 spectra are reported in the article is that OM2-D3 is better at describing intramolecular rearrangements such as the McLafferty reaction, see also original QCEIMS paper.[1]

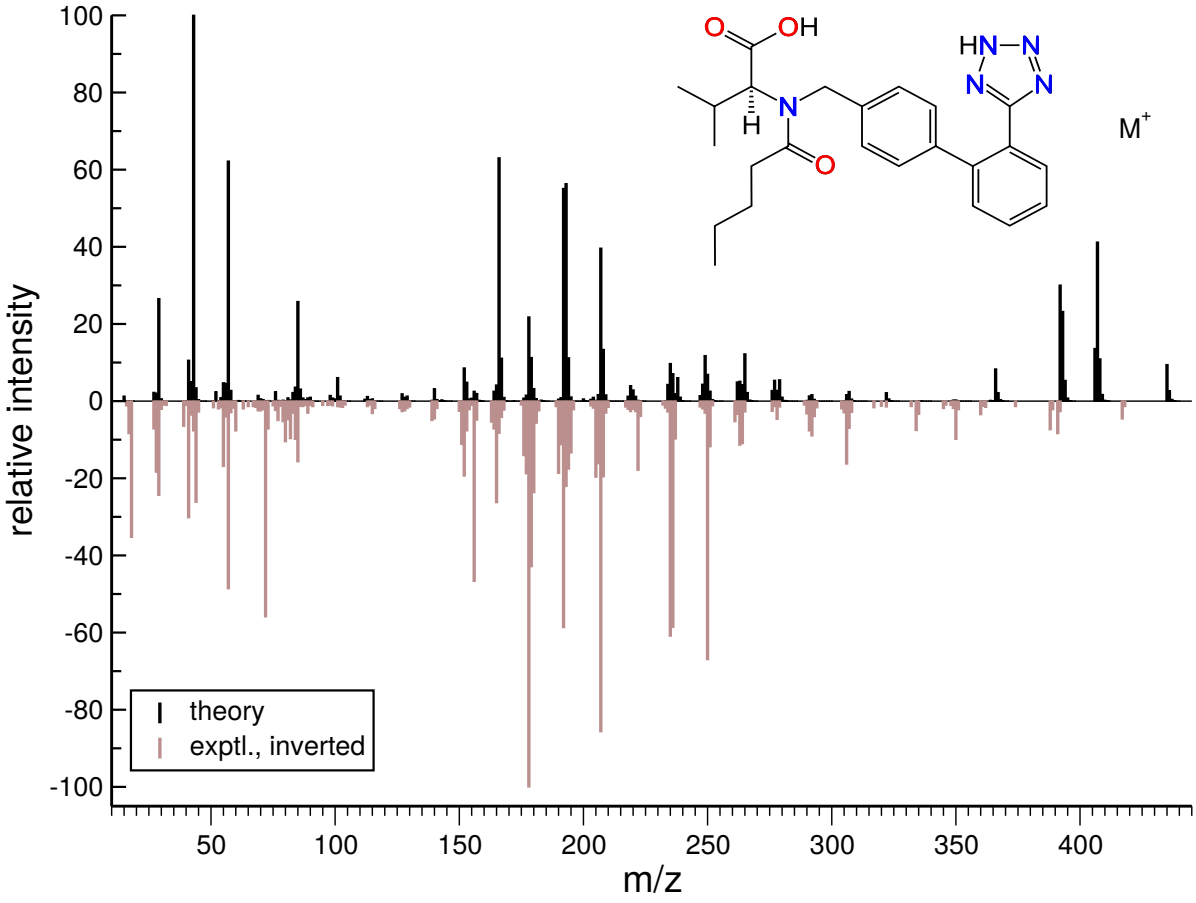

Figure 3: Calculated (DFTB3-D3) spectrum of **1** in comparison with the experimental spectrum.

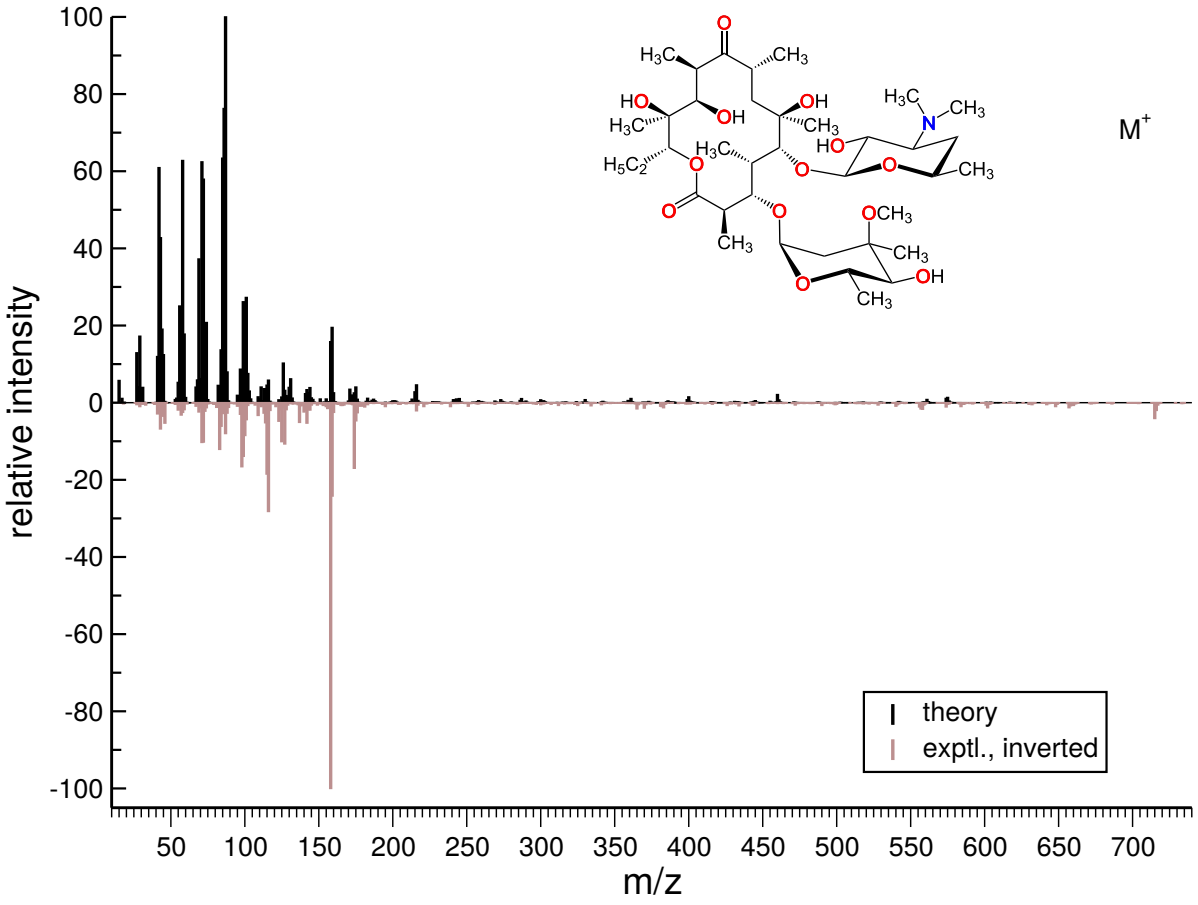

Figure 4: Calculated (DFTB3-D3) spectrum of **2** in comparison with the experimental spectrum. Note that the molecular ion (m/z=734) gives only a very weak signal in the experimental spectrum and none in the computed.

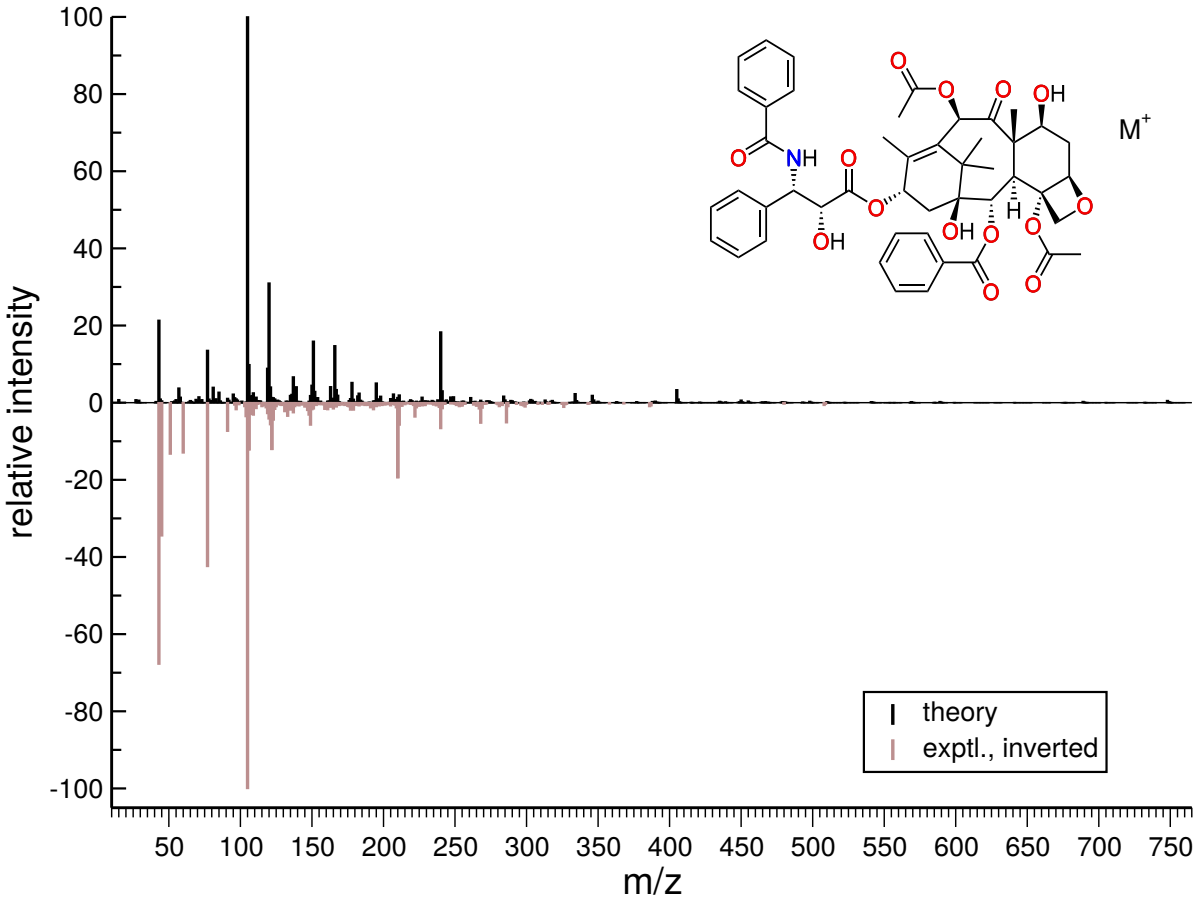

Figure 5: Calculated (DFTB3-D3) spectrum of **3** in comparison with the experimental spectrum. Note that the molecular ion (m/z=854) does not appear in either spectrum.

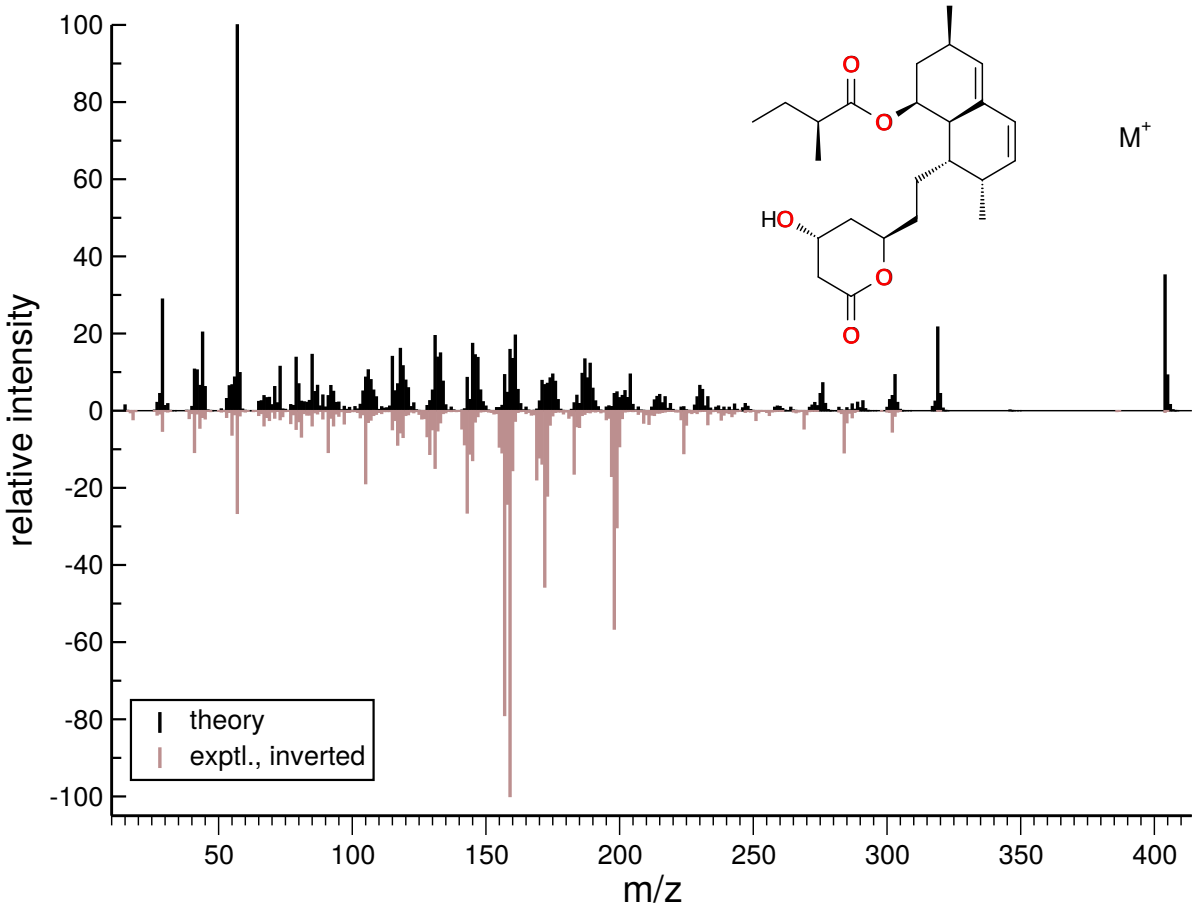

Figure 6: Calculated (DFTB3-D3) spectrum of **4** in comparison with the experimental spectrum.

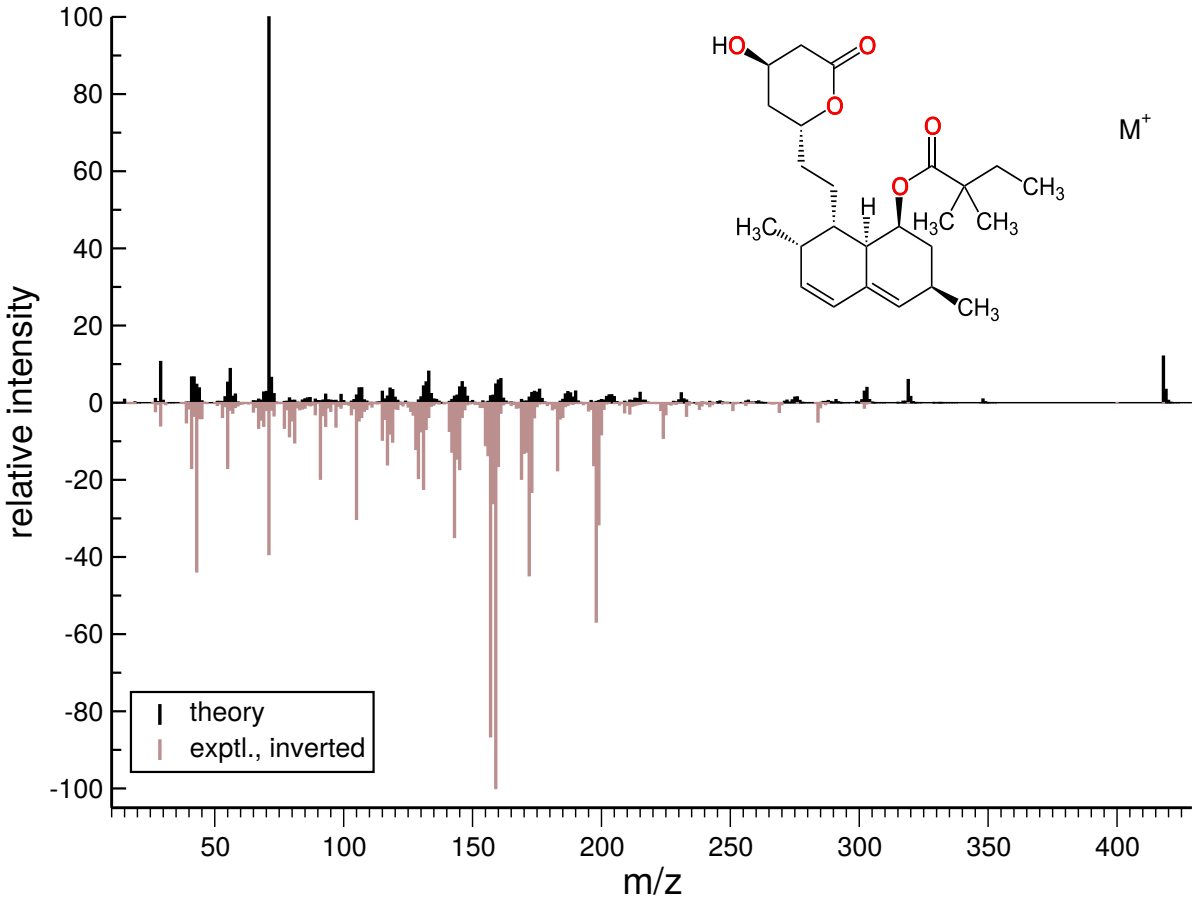

Figure 7: Calculated (DFTB3-D3) spectrum of **5** in comparison with the experimental spectrum.

References

(1) Grimme, S. *Angew. Chem. Int. Ed.* **2013**, *52*, 6306–6312.

(2) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics*; Cambridge University Press, 2009.

(3) Weber, W.; Thiel, W. *Theor. Chem. Acc.* **2000**, *103*, 495–506.

(4) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

(5) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.

(6) Mermin, N. D. *Phys. Rev.* **1964**, *134*, A112–A125.

(7) Warren, R. W.; Dunlap, B. I. *Chem. Phys. Lett.* **1996**, *262*, 384–392.

(8) Rabuck, A. D.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 695–700.

(9) MNDO2005, Version 7.0, W. Thiel, MPI für Kohlenforschung, Mülheim, Germany.

(10) (a) Aradi, B.; Hourahine, B.; Frauenheim, T. *J. Phys. Chem. A* **2007**, *111*, 5678; (b) T. Frauenheim, DFTB+ (Density Functional based Tight Binding) 2008, see http://www.dftb.org/.