Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Surface Science 223 (2004) 168–174
applied surface science
www.elsevier.com/locate/apsusc

Application of a genetic algorithm and a neural network for the discovery and optimization of new solid catalytic materials

U. Rodemerck, M. Baerns*, M. Holena, D. Wolf[1]

Institute for Applied Chemistry Berlin-Adlershof, Richard-Willstätter-Strasse 12, Berlin D-12489, Germany

Abstract

In the process of discovering new catalytic compositions by combinatorial methods in heterogeneous catalysis usually various potential catalytic compounds have to be prepared and tested. To decrease the number of necessary experiments an optimization algorithm based on a genetic algorithm for deriving subsequent generations from the performance of the members of the preceding generation is described. This procedure is supplemented by using an artificial neural network for establishing relationships between catalyst compositions—or more general speaking—materials properties and their catalytic performance. By combining a trained neural network with the genetic algorithm software virtually computer experiments were done aiming at adjusting the control parameters of the optimization algorithm to the special requirement of catalyst development. The approach is illustrated by the search for new catalytic compositions for the oxidative dehydrogenation of propane.

Keywords: Genetic algorithm; Neural network; Catalytic materials

1. Introduction

In the last 5 years combinatorial and high-throughput methods which were initially developed for drug discovery have been widely applied in the search for new inorganic materials for the use in heterogeneous catalysis [1–14]. Since industrial catalysts often consist of up to 10 chemical elements out of more than 50 chemical elements which principally can serve either as catalytically active materials or catalyst modifiers, a very vast parameter space has to be searched. Taking into consideration the fact that also the concentration of the individual component plays a major role for establishing the inorganic structure of the catalysts the number of possible catalytic active compositions is much higher than ever can be prepared and tested even applying high-throughput methods which can handle thousands of samples per week [15]. This fact makes it necessary to use intelligent methods for designing the experiments, i.e. restrict the experiments to a number which can be handled in a suitable period but still maintaining the chance of finding a catalyst having the desired activity, selectivity and stability.

To reach this goal the researcher can follow different paths: The way also used in non-combinatorial catalysis is to take into consideration all the knowledge which is already known about the targeted reaction and about known catalysts and combine this knowledge to derive a hypothesis on catalyst functions which are necessary for catalyzing the target reaction and also on catalyst functions to be avoided since they will catalyze

* Corresponding author.
E-mail address: baerns@aca-berlin.de (M. Baerns).
[1] Present address: Degussa AG, Projekthaus Katalyse, Industriepark Hoechst Geb. G830, D-65926, Frankfurt am Main, Germany.

0169-4332/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0169-4332(03)00919-X

undesired side reactions (rational design of catalysts). Another way is to use mathematical optimization routines which are able to find the global maximum of the function describing the dependence between catalytic reactivity and composition of the inorganic material. Several optimization methods have been suggested for this purpose and first promising results have been obtained in applying such methods [16–19].

In the following, a methodology using both approaches is described for the search for new catalytic active and selective materials for the oxidative dehydrogenation of propane to propene. A rational choice of potentially active catalyst components is used for defining a starting library of catalysts which is then further optimized by applying an evolutionary strategy using genetic algorithms (GA) for creating new generations of catalytic materials. High-throughput methods are applied for time-saving preparation and testing of the materials in the optimization loop going from the first catalyst library to the optimal catalyst composition in the parameter space to be searched.

Artificial neural networks (ANN) are used for modeling the whole parameter space and for identifying compositional areas where high propene yields can be expected. Moreover, a trained ANN was connected to the optimization software to make virtual simulation of catalyst performance on the computer possible.

2. Application of a genetic algorithm for searching the optimal catalyst composition

The application of a GA for searching the optimal composition of heterogeneous catalysts has been initially described by Wolf and co-workers [17,20]. Therefore, only a short description of the principles is given here. The GA applied in catalyst optimization is—as in biological systems—based on the operators ''mutation'' and ''crossover'' as well as selection of the fittest individuals. Applied to catalyst composition new catalysts are created from the best catalysts of the former generation by qualitative mutation (one or more chemical elements or compounds are replaced by others), quantitative mutation (the concentration of one or more compounds in the inorganic material is changed) and crossover (components are exchanged between two materials of the former generation generating two new materials). The ''survival of the fittest'' is reached by using a ranking selection, which gives a higher probability of selection to catalysts with better rank after high-throughput testing of a generation of catalysts.

Several applications to different types of catalysts (mixed metal oxides, supported metal catalysts) and reactions (oxidative dehydrogenation of ethane and propane, total oxidation of hydrocarbons, water gas shift reaction, light paraffin isomerization) have shown that this optimization routine is able to find the optimal catalyst composition within a limited number of experiments (100–1000) [13,14,17,19,21–25].

The first stage in searching new catalysts for the oxidative dehydrogenation of propane is to establish a pool of elements, which are then randomly combined to mixed metal oxide catalysts to generate a first generation of mixed metal oxides. From the possible reaction mechanisms oxides with the required properties were chosen (suitable redox properties: V, Mo, Mn, Fe, Ga; ability to dissociative adsorption of oxygen: La; non-removable lattice oxygen: B, Mg; see [14]). Fifty-six catalyst compositions each consisting of four of the above elements are generated randomly to establish the first generation. The composition of these catalysts has been described elsewhere [17]. Fig. 2 shows that the eight elements are nearly equally distributed over the population of the first generation. After high-throughput preparation and testing of the catalyst samples of the first generation the second generation was generated by applying the operators mutation and crossover and also tested for its catalytic performance. The probabilities for applying the evolutionary operators is changed during the optimization in a self-adapting way as it was described in [17]. The control parameters A and B are set to 0.5 (for the meaning of these parameters see [17], p. 67). Table 1 displays the change of the probabilities of crossover, qualitative and quantitative mutation during the optimization.

Table 1
Probabilities of applying the operators crossover (P cross) qualitative mutation (P mut) and quantitative mutation (P quant)

Generation	Y best	Y mean	P cross	P mut	P quant
1	7.5	2.8	0.41	0.50	0.09
2	7.7	4.0	0.37	0.50	0.13
3	9.0	4.9	0.36	0.50	0.14
4	9.0	5.9	0.34	0.50	0.16
5	9.1	5.9	0.34	0.50	0.16

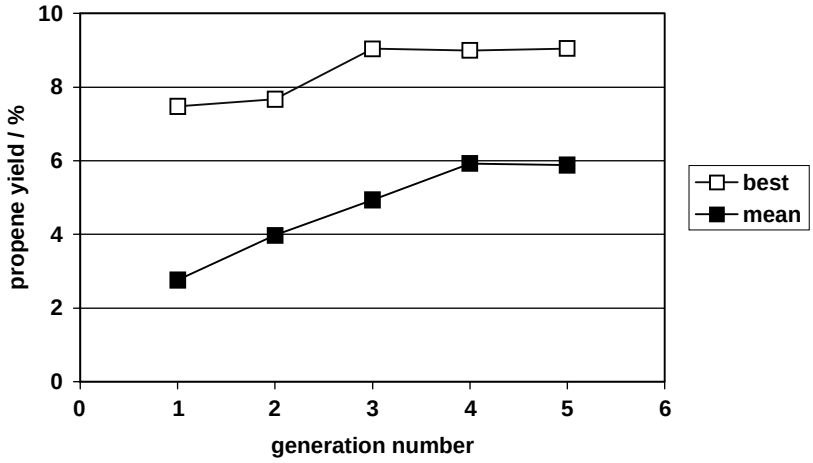

Fig. 1. Increasing catalyst performance during GA run: propene yields of best catalyst created in each generation (□) and mean propene yields over all catalysts of each generation (■).

The optimization of catalysts is illustrated in Fig. 1 for five generations showing the increasing yield of the selective product propene when going from one generation of catalytic materials to the next one. After some generations the optimization routine concentrates on catalyst compositions containing Mo, Mg, Ga and V whereas the other elements play a minor role (Fig. 2) [26]. The fact that after some generations the best catalysts are composed of only some elements clearly indicates that the algorithm focuses on an optimal catalyst composition.

3. Artificial neural networks—a valuable tool in combinatorial catalysis

In the last decade ANNs have been widely used in catalysis for modeling the relationships between

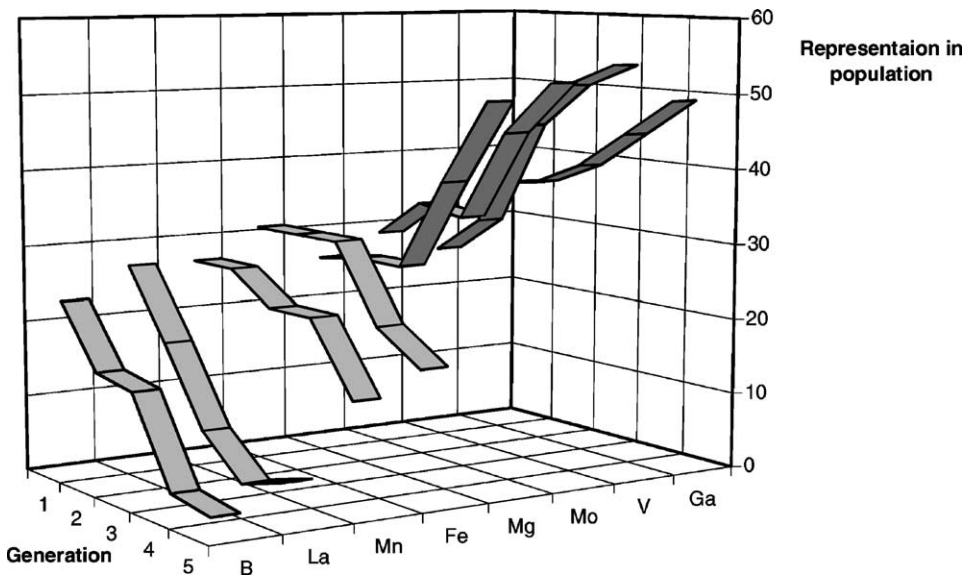

Fig. 2. Changing of catalyst composition during GA run.

catalyst composition and the catalytic performance [27–32]. ANNs have the advantage that no previous knowledge about the mathematical function describing these relationships must be known. Therefore, ANNs are well suitable for modeling the multi-parameter space searched in combinatorial catalysis for new active and selective materials.

In a very recent publication we reported on the training of 230 different ANN architectures with 211 data sets which were derived during catalyst optimization for the oxidative dehydrogenation of propane [33]. The best suited trained net was used to predict material compositions that are expected to give high propene yields. The experimental validation of the predicted results shows that the ANN is very well suited to predict the performance of catalyst composition which were not contained in the pool of data on which training was based (Fig. 3).

Moreover, we have very recently shown how knowledge in the form of rules can be extracted from a trained ANN [33,34]. In the following we present two examples for such rules describing the relationship between catalyst composition and performance for the oxidative dehydrogenation of propane:

- *Rule 1*: Propene yields are above 8% *if* the proportion of Ga is between 24 and 33% *and* the proportion of Mg is between 31 and 39% *and* the proportion of Mo is between 0 and 7% *and* remaining proportions are 0%.

Table 2
Experimental validation of rule 1 (examples 1–3) and rule 2 (examples 4–6) extracted from a trained ANN

Example	Composition (mol%)				Propene yield (%)	
	Ga	Mg	Mo	V	Predicted	Experimental
1	32	32	7	29	8.1	8.2
2	27	36	6	31	8.1	8.4
3	32	33	5	30	8.3	8.0
4	38	31	8	23	8.3	7.9
5	38	31	9	22	8.4	8.3
6	38	32	9	21	8.4	8.2

- *Rule 2*: Propene yields are above 8% *if* the proportion of Ga is approximately 38% *and* the proportion of Fe is between 0 and 12% *and* the proportion of Mg is between 29 and 36% *and* the proportion of Mo is between 0 and 9% *and* remaining proportions are 0%.

These rules have been validated by preparation and testing of catalyst samples fulfilling the prerequisites described it. Table 2 shows that the experimentally determined propene yields meet the predicted values very well.

4. Connecting GA and ANN

In the following an idea is presented which uses a trained ANN as a simulator of virtual experiments to

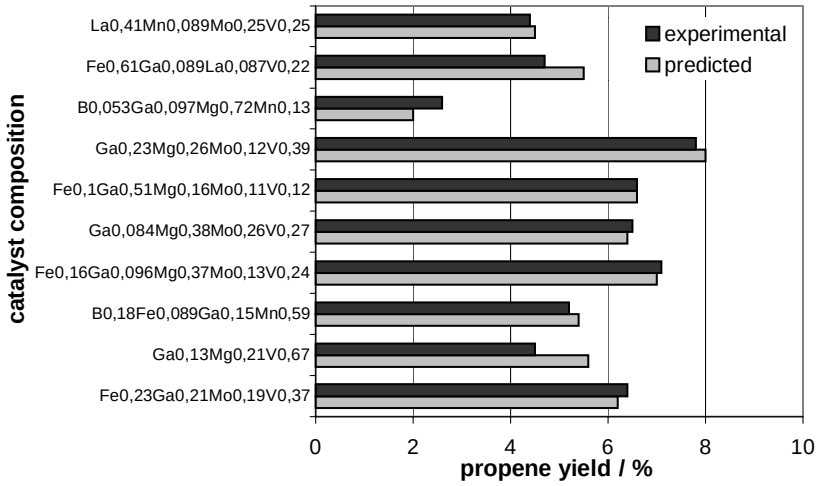

Fig. 3. Experimental validation of catalysts predicted by the trained ANN.

study the influence of the various parameters of an GA optimization routine, i.e. the proportions of crossover, mutations and members per generation on its performance.

Up to now little is known about the topology of the parameter space describing the relationship between catalyst composition and their performance for the different types of catalytic reactions. Therefore, it seems to be impossible to tune an optimization method like the GA to the requirements of the reaction under study. However, if enough data for one reaction are available, modeling of the parameter space by training an ANN can be used to get some knowledge about the topology from that model. Moreover, the trained ANN can be used as a source of virtual experimental data for studying the influence of the control parameter, i.e. the proportion of and type of mutation and crossover and the population size of the individual generations of catalytic materials on the speed of convergence to the global optimum of the catalyst composition. Due to using computer programs a high number of different sets of control parameters can be tested in a reasonable time.

An ANN (architecture 7-5-1, see [33]) trained with 328 data points for the oxidative dehydrogenation of propane to propene has been used to generate the virtual experimental data, i.e. to predict propene yields for given compositions of the catalytic materials. To determine the global optimum (maximum of propene yield) sequential-quadratic-programming modification of the Levenberg-Marquardt method [35] was applied. A maximum propene yield of 9.1% was determined for the catalyst composition $Ga_{0.21}Mg_{0.40}Mn_{0.21}V_{0.18}O_x$.

For studying the different sets of parameters of the GA a software was written which connects the trained ANN with the GA software. Most importantly, the influence of the population size (number of materials in one generation) was studied by computing three runs of the GA of 50 generations each with generation sizes of 28, 56, and 280, respectively, i.e. a high number of catalyst compositions and their performance were studied in virtual computer experiments. The results are depicted in Figs. 4 and 5. Fig. 4 shows that the population size does not play a major role in the speed of convergence. In all cases after five generations a 90% value of the global maximum of the propene yield is reached; this confirms that the GA is able to find the global optimum within a limited number of experiments. More experiments (6th to 50th generation) result only in a marginal improvement of the catalyst; this is a very important result as it can help to save experimental effort. Fig. 5 displays the compositions of the

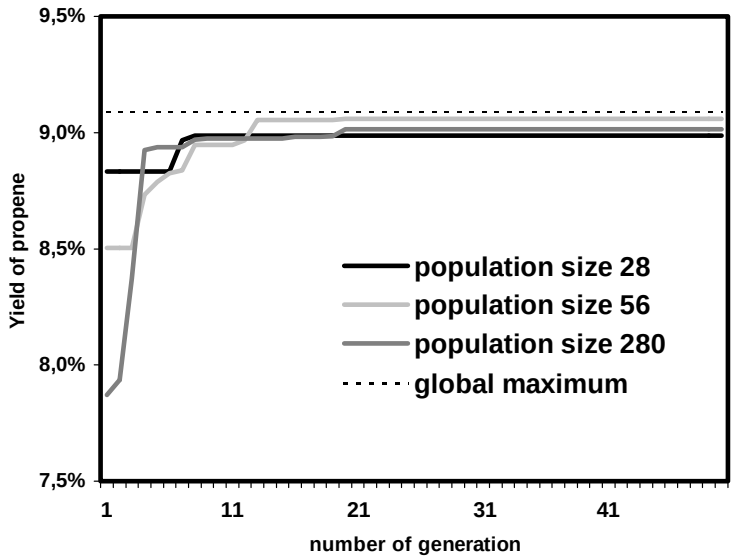

Fig. 4. Influence of the size of population on the speed of convergence of the GA.

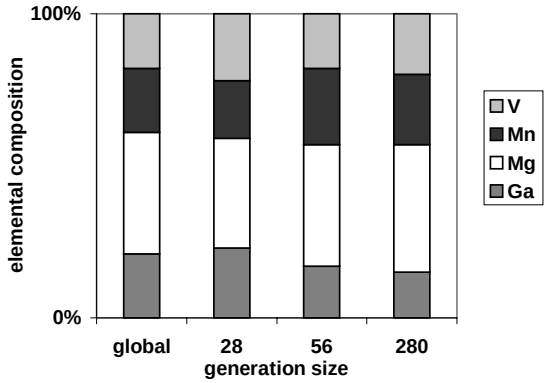

Fig. 5. Composition of the optimal composed catalyst found by the four runs of GA applying different population sizes compared to the composition of the optimal composed catalyst.

optimal composed catalytic materials and the catalysts optimized by the three virtual runs of the GA. From a chemical point of view the composition of the mixed metal oxides is nearly the same; the slight differences are not expected to influence the structure of the inorganic material.

5. Conclusions and outlook

Genetic algorithms are well suited in the field of combinatorial catalysis to discover optimal catalyst compositions within a limited number of experiments. However, high-throughput experimentation for preparing and testing the materials samples is a prerequisite to run the necessary experiments to shorten the time of the development.

ANNs are well suited to model the relationship between catalysts composition and their reactivity in the vast parameter spaces under investigation in combinatorial catalysis. Even if nothing is known about the mathematical functions describing the dependencies of the parameters an ANN can be used to extract knowledge from the data pool generated by high-throughput experiments. Moreover, rules can be extracted from a trained ANN which give the chemist a hint on the catalyst components influencing its performance; on a more fundamental level, it is expected that the scientist is inspired to create hypotheses on the function of the individual components in the reaction mechanism.

Virtual computer experiments can help to adjust the parameters controlling the speed of convergence of the optimization routine of a GA to the special requirements of the catalytic reaction under investigation. It was shown that—from a high-throughput experimentation point of view—only a limited number of experiments is necessary to find materials having nearly the composition and performance of the optimal catalyst in the vast parameter space under investigation. Further studies on the influence of the kind and proportion of qualitative and quantitative mutation as well as crossover are under preparation and will be published soon [36].

Acknowledgements

Financial support by the Federal Ministry of Education and Research of Germany and by the State of Berlin is gratefully acknowledged. The authors thank Mrs. Divia Kumar for programming and Mrs. Renate Jentzsch for experimental assistance.

References

[1] P. Cong, R.D. Doolen, Q. Fan, D.M. Giaquinta, S. Guan, E.W. McFarland, D.M. Poojary, K. Self, H.W. Turner, W.H. Weinberg, Angew. Chem. 111 (1999) 867.
[2] Y. Liu, P. Cong, R.D. Doolen, H.W. Turner, W.H. Turner, Catal. Today 61 (2000) 87.
[3] S.M. Senkan, S. Ozturk, Angew. Chem. 111 (1999) 867.
[4] K. Krantz, S. Ozkru, S. Senkan, Catal. Today 62 (2000) 281.
[5] M. Richter, M. Langpape, S. Kolf, G. Grubert, R. Eckelt, J. Radnik, M. Schneider, M.-M. Pohl, R. Fricke, Appl. Catal. B: Environ. 36 (2002) 261.
[6] D.E. Akporiaye, I.M. Dahl, A. Karlsson, R. Wendelbo, Angew. Chemie 110 (1998) 629.
[7] D. Akporiaye, I. Dahl, A. Karlsson, M. Plassen, R. Wendelbo, D.S. Bem, R.W. Broach, G.J. Lewis, M. Miller, J. Moscoso, Microporous Mesoporous Mater. 48 (2001) 367.
[8] J. Klein, C.W. Lehmann, H.-W. Schmidt, W.F. Maier, Angew. Chemie 110 (1998) 3557.
[9] A. Wolf, F. Schüth, Appl. Catal. A: Gen. 226 (2002) 1.
[10] I. Hahndorf, O. Buyevskaya, M. Langpape, G. Grubert, S. Kolf, E. Guillon, M. Baerns, Chem. Eng. J. 89 (2002) 119.
[11] C. Hoffmann, H.-W. Schmidt, F. Schüth, J. Catal. 198 (2001) 348.
[12] C. Lettmann, H. Hinrichs, W.F. Maier, Angew. Chemie 113 (2001) 3258.
[13] U. Rodemerck, D. Wolf, O.V. Buyevskaya, P. Claus, S. Senkan, M. Baerns, Chem. Eng. J. 82 (2001) 3.

[14] O.V. Buyevskaya, A. Brückner, E.V. Kondratenko, D. Wolf, M. Baerns, Catal. Today 67 (2001) 369.
[15] S. Senkan, Angew. Chem. 113 (2001) 322.
[16] A. Holzwarth, P. Denton, H. Zanthoff, C. Mirodatos, Catal. Today 67 (2001) 309.
[17] D. Wolf, O.V. Buyevskaya, M. Baerns, Appl. Catal. A: Gen. 200 (2000) 63.
[18] J.N. Cawse, Experimental Design for Combinatorial and High Throughput Materials Development, Wiley, Hoboken, New Jersey, 2003.
[19] A. Corma, J.M. Serra, A. Chica, in: E.G. Derouane, V. Parmon, F. Lemos, F.R. Ribeiro (Eds.), Principles and Methods for Accelerated Catalyst Design and Testing, Kluwer, Dordrecht, 2002, p. 153.
[20] D. Wolf, M. Baerns, in: J.N. Cawse (Ed.), Experimental Design for Combinatorial and High Throughput Materials Development, Wiley, Hoboken, New Jersey, 2003, p. 147.
[21] G. Grubert, D. Wolf, N. Dropka, S. Kolf, M. Baerns, Rapid discovery of new catalytic materials for the oxidative dehydrogenation of ethane to ethylene by an evolutionary approach, in: Proceedings of the Fourth World Congress on Oxidation Catalysis, Berlin/Potsdam, Germany, 2001, p. 113.
[22] M. Baerns, O. Buyevskaya, G. Grubert, U. Rodemerck, in: E.G. Derouane, V. Parmon, F. Lemos, F.R. Ribeiro (Eds.), Principles and Methods for Accelerated Catalyst Design and Testing, Kluwer, Dodrecht, 2002, p. 85.
[23] G. Grubert, E. Kondratenko, S. Kolf, M. Baerns, P.v. Geem, R. Parton, Catal. Today 81 (2003) 337.
[24] O.V. Buyevskaya, D. Wolf, M. Baerns, Catal. Today 62 (2000) 91.
[25] M. Langpape, G. Grubert, D. Wolf, M. Baerns, DGMK-Tagungsbericht 4 (2001) 227.
[26] J.N. Cawse, M. Baerns, personal communication.
[27] T. Hattori, S. Kito, et al., Catal. Today 23 (1995) 347.
[28] S. Kito, T. Hattori, Y. Murakami, et al., Appl. Catal. A: Gen. 114 (1994) L173.
[29] M. Sasaki, H. Hamada, Y. Kintaichi, T. Ito, et al., Appl. Catal. A: Gen. 132 (1995) 261.
[30] Z.-Y. Hou, Q. Dai, X.-Q. Wu, G.-T. Chen, et al., Appl. Catal. A: Gen. 161 (1997) 183.
[31] B.K. Sharma, M.P. Sharma, S.K. Roy, S. Kumar, S.B. Tendulkar, S.S. Tambe, B.D. Kulkarni, et al., Fuel 77 (1998) 1763.
[32] K. Huang, F.-Q. Chen, D.-W. Lü, et al., Appl. Catal. A: Gen. 219 (2001) 61.
[33] M. Holena, M. Baerns, in: J.N. Cawse (Ed.), Experimental Design for Combinatorial and High Throughput Materials Development, Wiley, Hoboken, New Jersey, 2003, p. 163.
[34] M. Holena, M. Baerns, Catal. Today 81 (2003) 485.
[35] K. Schittkowski, Ann. Operations Res. 5 (1985) 485.
[36] M. Holena, M. Baerns, in preparation.