

汑 𣎵 𣎶 𣎷 汑 汑 𣎵 𣎶 𣎷 𣎵 汑 𣎶 𣎷 汑

Twenty Characters

汑 𠂔 𠂔 𠂔 汑 汑
 𠂔 𠂔 𠂔 𠂔 𠂔 𠂔 𠂔 𠂔

廿百卅先廿沽 卅拾十侯十卅廿卅侯廿

Kristian Rother

Twenty Characters

Kristian Rother

Credo

Where the established models fail, is the outer frontier of science. There, one has to seek new models that are honest, simple, and intuitively understandable. Models that are beautiful.

From „Missing Link“ to „Twenty Characters“

In the summer of 2004, an experimental get-together took place under the title of „Missing Link. Art meets Biomedicine – Public Understanding of Art and Science“: Artists and scientists met in the countryside close to Berlin in order to talk about their fields of expertise, their technical terminologies, and their work methods. The title „Missing Link“ was understood as a hint as much at a lack of connections between the artistic and scientific disciplines as at the possibility of discovering (or creating) points of contact.

Traditionally, artists in the sciences have been working as visualisation specialists: The most well-known example may be *the* popular icon of scientific progress in the 20th century, the representation of the double-helix based on the model by artist Odile Crick. This kind of cooperation has practically come to an end by now. The same goes for the other side: After the ideal of realist depiction lost its predominance in the arts, the custom of contracting physicians to hold lectures on anatomy or even dissections at the art academies was abandoned as well. So are there any fields of cooperation, are there any common features left, and where could these be found?

In public discussion, the sciences are often seen as dealing with „reality“ and „objective facts“, while the arts are suspected of potentially being off with the fairies. Interestingly, this popular dichotomy never came up during the Missing Link meeting and it looks as if it can be relegated to the feuilleton. The highly specialized scientists were very conscious of their images' model character and all the artists dealt with reality (and realities), ranging from strictly observing figurative painting to artistic-sociological interventions.

Common experiences were mainly linked to the processes of creative work and could be summarized as „work on the invisible“. One of the meeting's result was a show at the Berlin Museum of Medical History in 2005¹. A second project which has been partly inspired by the „Missing Link“ meeting is „Twenty Characters“, which in its complex linking of image and writing, of striking visuality and practicality manages to accomplish a real crossover between disciplines: The attempt to reinvest an abstract type referring to invisible structures with a sense of haptic connection to the signified shape - a „Found Link“?

Käthe Wenzel

1 Wolfgang Knapp, Käthe Wenzel, Cornelius Frömmel, Thomas Schnalke. Missing Link: Kunst trifft Biomedizin. Katalog zur Ausstellung. Berlin 2005.

Summary

This manuscript proposes a novel and more intuitive way to notate proteins, inspired from Chinese calligraphy. The rule set by which the characters of the Chinese language are composed is introduced. By these rules, semantic, phonetic and aesthetic content can be presented in a highly condensed form. Here, the Hanzi Transformation was developed, a method capable of converting any picture into the context of asian calligraphy. In a series of calligraphic studies, abstract figures of persons and objects approximating the Chinese character form were created. This stretched the aesthetics of the calligraphic painting technique into a completely new context, and, in the most successful cases, it allowed for the construction of new character symbols still retaining the pictographic information. Consequently applying the Hanzi Tranformation method to biomolecular structures, a series of proteins could be depicted as well. Creating these protein calligraphies requires to immerse into their three-dimensional structures, and it facilitated understanding their architecture considerably. In the final part of this study, the primary structure was transformed as well: The chemical structures of the amino acids were used to construct a set of 20 characters as an alternative one-letter-code. All important properties, like charges, aromaticity and chain size are directly visible in the symbols. Thretherefore, this character set allows to perceive similarity on a chemical level much easier than the latin characters commonly used. The font created here is aimed to circumvent limitations in human perception that are an obstacle in finding information in multiple protein sequence alignments.

Acknowledgements

This book would not have been realized without the photographs and ongoing support by Magdalena Musielak, the advice on art by Käthe Wenzel. I thank the voluntary reviews from Marta Kubala, Kara van Aelst, Goran Martinovic, Hans-Christian Schütz, Joost Klüßendorf, Cristian Hoof, and Markus Rother, and Andrea Schmidt for the exciting dialogue about typography. I'd like to acknowledge Manuala Jacob, Ulrike Bley, Stefan Günther and Uwe Müller for providing protein images. I owe respect to my models Tanja Fromm, Eva-Maria Keding, Maria Luz, Jasmin Strackharn, Sabrina Hofmann, Dirk Weinrich, Christoph Gille, Ralf Hofestätter and several unknown people, to Nils Goldmann for asking me why it is important to save strokes, and to Marei Gatemann for taking care of me immediately after my accident. Finally, I'd like to thank the two persons who initiated my work: Bingyu Zhu, who gave me the brush, and Cornelius Frömmel who taught me that science and art are inseparable.

汁 百 寺 先 汁 沽 升 格 十 侯 十 升 汁 寺 侯 汁

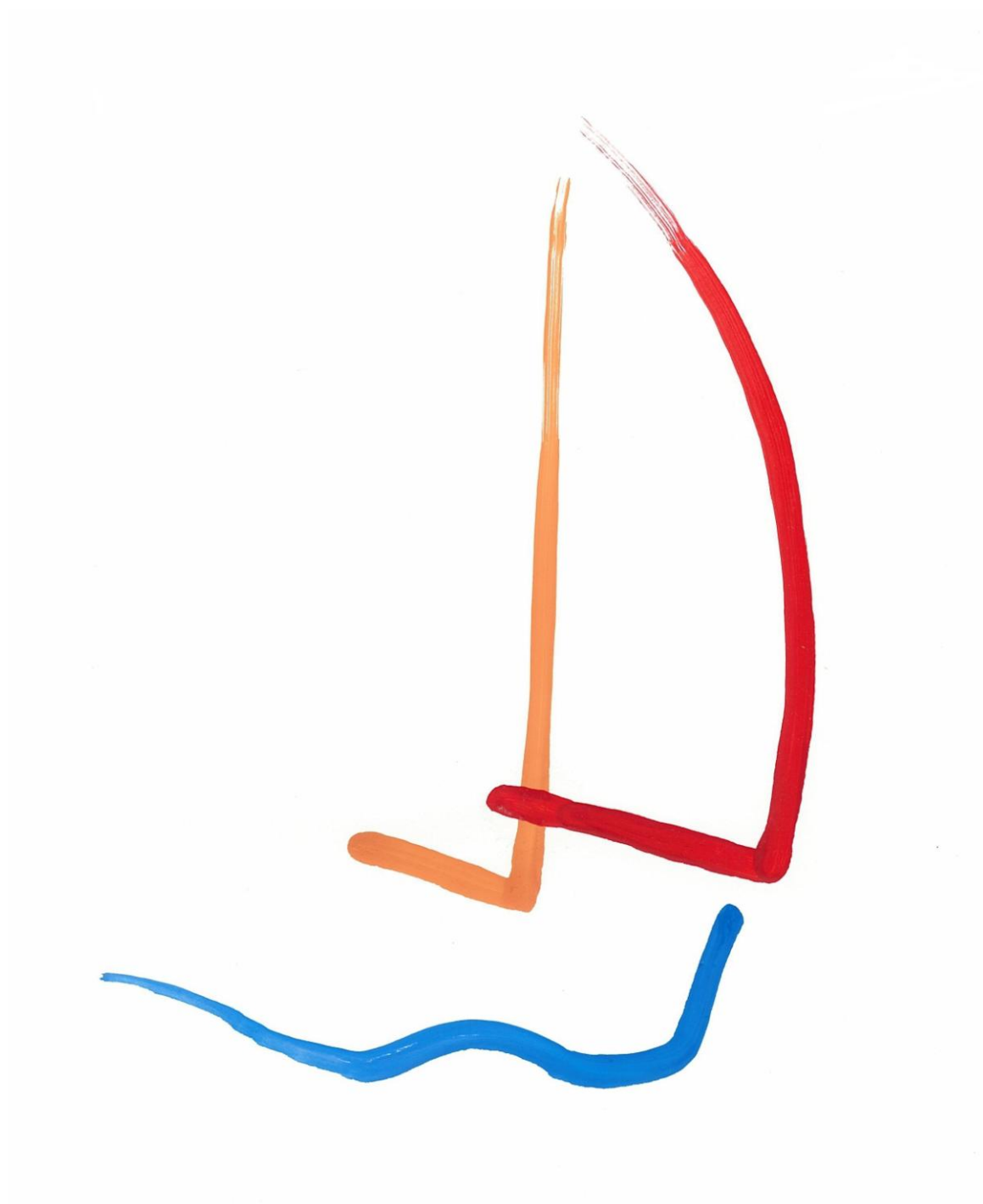


Figure 1: "Backbord", guache 2005.

汁百吉先汁沽 升格十僕十升汁吉僕汁

The brush strokes in all calligraphy show to the observer the path of the calligraphers hand. They are never static, straight constructs. A stroke can convey force (*Li*) that arises e.g. from the optical tension in a line with reinforced ends (see Figure “Shui Luo Shi Chu”, to the right). It can also carry dynamic (*Qi*), when the end of one stroke points to another component of the picture (see Figure 17). Ideally, one character should be completed within a single breath, without setting the brush aside.

In the composition of a picture, there are usually lots of empty space. It hosts a few large and several small components, that form an asymmetric equilibrium (see Figure 16). Structural components are often similar without repeating themselves. The well-balanced combination of these features creates *Chi Yuan* - the living stroke (Kwo 1990).

Brush Strokes

The more than 50,000 known Chinese characters (*Han Zi*) reflect the culture, history and mentality of their developers. Unlike the latin alphabet, many characters originating from pictographs can still be recognized (see Figures 6 and 7). But for many others, abstraction has impaired the pictographic and contextual information, and they are mere symbols. How is it then possible to learn such a vast language? Fortunately, the entire alphabet is built by a hierachical logic: Every character is built from a combination of 30 basic strokes. These are composed to simple shapes or *radicals*. The radicals already convey semantic (*sun, tree, walk*) or phonetic (*chang, lei, mu*) information. One or a few radicals compose a character. This way, the enormous number of combinations of the basic strokes can be subject to understanding. An analogous scheme how the human mind can grasp such almost infinite complexity has been formulated for music and protein sequences (Frömmel 2000).

Due to the high number of characters that stand for entire words or concepts, Chinese text is compact. The semantic information carried by a single character is very high. And unlike languages written in latin or related character systems, Han Zi contain only inaccurate phonetic information – in fact, several languages exist for the same script. As a consequence for people learning Mandarin this means that small mistakes and bad pronunciation can have devastating consequences – like me accidentally ordering two pounds of dragon dicks at a fruit merchant.

The compact nature of Hanzi characters can be expressed in numbers as well: Behr et al. (Behr 2002) compared the size and compressibility of the Bible and the UN Treaty Collection in different languages. The compressed versions of these texts using the PPM algorithm (Cleary 1984) have about the same size, indicating that the information in these texts is about the same. But, the uncompressed Chinese texts have only 22.9% of the number of characters as the English ones. To store a Chinese character 6.1 bit were necessary, while only 1.62 bit per English character were used¹. This means that English

1 These numbers were calculated from the online version of the paper, dividing the total number of characters by the compressed byte size*8. Note that the number of characters for the GB2312-encoded text is half the file size.

汑 𠂇 𠂇 𠂇 汑 汑 𠂇 𠂇 𠂇 𠂇 汑 汑 𠂇 𠂇 汑

could be written using the four DNA bases with the same number of characters, if the words would be redefined properly.

After recognizing, that I had intuitively reproduced the historical development of abstracting images while painting the sailing boat, I decided to create pictographs in a similar way systematically.



Figure 2: "Reading in the subway line 1", Chinese ink, 2006.

II. Hanzi Transformation

Most of the sketches I drew every now and then allowed to clearly recognize the person or object, but in itself they were not very interesting. The opposite was the case for the *Han Zi*, the Chinese characters. I found both extremes artistically unsatisfactory, and strived to combine both to emphasize the important things on my sketches more. By proper choice from among the 30 elementary brush strokes, it should be possible to produce abstracted but still meaningful pictures. Consequently, the process by which the Han Zi alphabet had emerged should be reproducible to create new characters from that. Even more, the process could be optimized in such a way, that the resulting character and the original sketch are still alike.

汁 吉 吉 先 汁 沽 十 格 十 庚 十 十 汁 吉 庚 汁

The central hypothesis of this work is that a picture can be transformed by subsequent abstraction into a pictographic character according to the rules of written Mandarin in such a way, that the original motif is still recognizable.

Converting Sketches

To test this hypothesis, I have developed the method of Hanzi Transformation. The transformation started from a sketch drawn on-site by pencil, for which a suitable pictographical representation was developed later in the studio. Practically this meant sparing a lot of lines from the initial picture. Many of the sketches were drawn in short time or on moving subjects. This enforced already a reduction of details in the sketch (compare to Nicolaidis 1990).

The pictographs were constructed incrementally: first, the most characteristic component of the motif was searched and added to the pictograph, then the second most important, and so on. As soon as the motif was recognizable in the pictograph, this process was terminated.

In a second step, each of the found components was substituted by one of the basic 30 brush strokes. This was often possible, because the basic strokes are defined in their topology and direction, but may vary in length and thickness. In cases where no suitable basic stroke was found, a similar stroke was created by rotating or mirroring a basic one.

In a final stage, the order of strokes was chosen in a way to allow for a flowing movement of the brush. Also, individual strokes were emphasized or weakened to balance the composition. Some of the resulting calligraphies were shrunk into the square-shape that the original *Han Zi* characters have.

Transformed Motives

For the Hanzi Transformation I deliberately chose motifs for which no singular characters exist in Chinese dictionaries: The Brandenburg Gate, particular scientists, an Esso gas station. The transformation succeeded to different abstraction levels: Some motifs were transformed completely into characters (see Figure 19), others developed into minimalistic paintings similar to the sailing boat (see Figure 2 and 13). In some cases, I succeeded in creating a one-stroke calligraphy, where the boundaries between the strokes dissolve (see Figure 11 and 12). This result corresponds to the historical Xie Yi (*written ideas*) style. Generally, a broad spectrum of picture forms was created, most having relatives in asian art. By selectively replacing the calligraphy ink by primary colors, many of the paintings were added a strong decorative feature (see Figure 12 and 17).

After succeeding in this, it was only consequent to apply this technique to subjects from my everyday work.

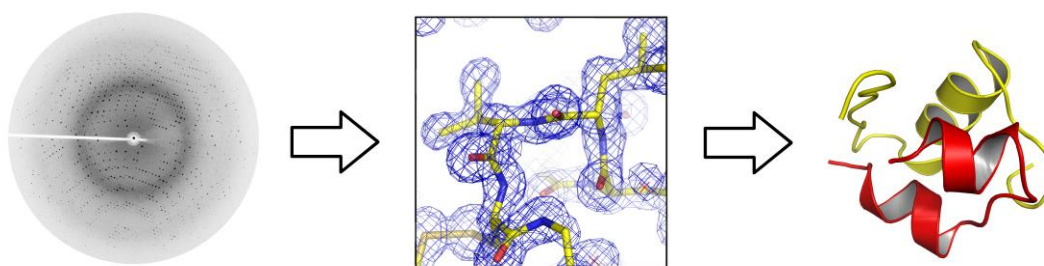


Figure 3: Left: Diffraction image of an insulin crystal in an x-ray beam. Recorded at the Synchrotron BESSY II in Berlin. Center: Electron density map of insulin (part). Right: three-dimensional structure model. Images provided by Uwe Müller.

III. Protein Structure Calligraphy

Proteins are, along with DNA and RNA, one of the central classes of molecules in living nature. They compose stable fibers, highly specific catalysts, molecule-size machines and many other things. This plethora of functions is achieved by each protein having a well-defined 3D structure. Analyzing the way in which proteins do their job is among the most exciting tasks in biochemistry.

The synthesis of proteins is a precisely defined assembly of components: Proteins are composed of 20 different building blocks, the amino acids. The amino acids differ in their size, shape and chemical properties. 40 to more than 1000 amino acids are assembled to protein chains in cells. Upon synthesis of the chain, the amino acids start interacting with each other. Some amino acids attract water, others prefer a water-free environment. Positive and negative charges on them attract each other, equal charges act repulsively. The resulting forces let the entire protein chain fold itself in space, until a stable, mostly globular shape has been reached. Therefore, the sequence² of all amino acids on the protein chains determines the structure and thereby the function of a protein².

Computer Images of Proteins

Many three-dimensional structures of proteins have been revealed by x-ray crystallography and NMR (nuclear magnetic resonance) experiments (see Figure 3). Being able to visualize the 3D shape of proteins has inspired biochemical research – and the researchers as well. In scientific journals the structure of protein molecules is depicted in many different ways: as arrangements of lines, tubes, arrows, as clouds of balls, and as objects that resemble heavily damaged potatoes. The choice depends on what structural features (like the topology of the protein chain, the shape of functional sites, or electrostatic forces) scientists want to emphasize.

To compensate for the loss of the third dimension in a printed image, a number of tricks

² The sequence of a protein is different from a DNA or genomic sequence, which consist of only four, and completely different building blocks.

has been invented: Stereo images have been used for a long time. Here, two adjacent pictures with slightly shifted perspective are prepared for the left and right eye. These images can be viewed in 3D, requiring some training - or a pair of stereo goggles that exist for exactly that purpose. With increasing computing power, high-resolution images of proteins have gained in popularity. Pseudorealistic light effects and depth cue can be rendered easily and in high quality. Practically any computer is able to generate such high-resolution images in real time, allowing a scientist to rotate even large structures and explore them. More complicated technologies like 3D glasses and 3D screens exist, but they are rarely seen. In contrast, animations have become increasingly popular – probably because they can be included in lectures and presentations. The technical barrier to create high-quality animations was high for a long time. Only recently, the eMovie plugin for the very popular PyMOL molecular viewer (DeLano 2006) provided an easy-to-use tool for creating animations at zero cost (Hodis 2007).

Despite all these technical support, the complexity of 3D structures is difficult to grasp. An observer easily loses orientation in the molecule, and gets lost in details. It is hard to find any resemblance from everyday experience. Moreover most proteins have no recognizable 'top' or 'bottom' side, while they can be rotated freely in space. A scientist browsing through a structure on a computer lacks an 'artificial horizon' that facilitates navigation.

Painting Protein Structures

Because proteins are chain-shaped molecules, they naturally contain a *Chi Yuan*-like path a brush could take. They are therefore almost ideal targets for the calligraphic technique. I undertook this experiment in order to discover, whether the abstraction helps to solve the orientation problem mentioned above. The Hanzi transformation method was adapted to produce more abstract one-line representations, containing a notion of secondary structural elements, instead of substituting parts by single strokes. For that, it was necessary to convert the three-dimensional topology of structural elements into a two-dimensional representation. In order to understand how a structure is built, I examined it from many different angles in a 3D viewer. Thus, I created landmarks for myself that helped me orienting. Finally, I painted the structures in a single brush stroke. The results were closer to the *Xie Yi* (written ideas) concept than the motifs converted before (see Figure 4). After the Hanzi Transformation, it was considerably easier to identify particular locations in a protein structure on sight. I recommend this approach to anybody who needs to make himself familiar with a particular protein.

For a peer review of my results, the protein calligraphies were shown to several protein structure scientists. In all cases where the „reviewers“ were familiar with a particular protein, they were able to identify it in the calligraphy almost at once. It is not unusual that scientists prefer drawn images of molecules over computer-generated ones – Branden & Tooze filled an entire book with them (Branden 1999). Hiring artists to draw protein structures was common practice before the high-quality computer graphics became common. Nowadays, drawn molecules are rarely seen in scientific publications, because

汁百吉先汁沽 汁谷十彙十汁汁吉彙汁

including the cue of the third dimension is difficult, and good artists are expensive. To compensate for this and justify its use, a drawn image needs to be considerably more illustrative, while being sufficiently accurate in its proportions.

The deduction from the structure of *Han Zi* characters to protein molecules should be obvious now. While this is appealing for the artist's eye, the scientist's mind is not satisfied yet. So far, the results are not useful for anything. In the next chapter, the outcome of the three previous ones will be merged to approach a practical problem.



Figure 4: Lysozyme (occurring in saliva and tears where it fights bacteria), Chinese ink, 2005.

汁百吉先汁沽 什谷十稊十什汁吉稊汁

transformed by simply changing the font in any word processor. This way, not only the recognition of common features is facilitated, the color feature is now free to be used for other means. I believe that this or any similar chemistry-based set of symbols have the potential to facilitate the analysis of protein sequences significantly, especially of multiple sequence alignments.

References

F.H. Behr, V. Fossum, M. Mitzenmacher, D. Xiao. 2002. Estimating and Comparing Entropy across Written Natural Languages using PPM Compression. TR-12-02. Harvard CS technical reports.

Carl-Ivar Branden, John Tooze. Introduction to Protein Structure. Garland 1999. ISBN 978-0815323044.

J.G. Cleary, and I.H. Witten, Data compression using adaptive coding and partial string matching. 1984. IEEE Transactions on Communications, 32 (4), 396-402.

Warren DeLano. 2006. Pymol (TM) Molecular Graphics System. www.pymol.org.

Werner Ebeling, Cornelius Frömmel. 1998. Entropy and predictability of information carriers. Biosystems 46(1-2):47-55.

Cornelius Frömmel. Das Flair der unendlichen Vielfalt. Humboldt Universität zu Berlin 2000.

Christoph Gille, Stephan Lorenzen, Elke Michalsky, Cornelius Frömmel. 2003. KISS for STRAP: user extensions for a protein alignment editor. Bioinformatics 19(18): 2489-2491.

Eran Hodis, Gideon Schreiber, Kristian Rother, Joel Sussman. 2007. eMovie: a storyboard-based tool for making molecular movies. Trends Biochem Sci 32(5):199-204.

L. Huang, A. Treisman A, H. Pashler. 2007. Characterizing the limits of human visual awareness. Science 317(5839):823-825.

Wolfgang Knapp, Käthe Wenzel, Cornelius Frömmel, Thomas Schnalke (Ed.): The Missing Link. Art meets Biomedicine. Public understanding of art and sciences. Berlin 2005. ISBN 3-89462-133.

Kimon Nicolaides. The Natural Way to Draw: A Working Plan for Art Study. Houghton Mifflin 1990. ISBN 978-0395530078.

Da-Wei Kwo. Chinese Brushwork in Calligraphy and Painting: Its History, Aesthetics, and Techniques. Dover 1990. ISBN 978-0486264813.

一	西	古	沽
十	汁	𠂇	𠂇
士	𠂇	𠂇	𠂇
主	𠂇	𠂇	𠂇
𠂇	𠂇	𠂇	𠂇

15

竹首节无节 叶谷十横十叶竹节横叶



Figure 6: Bamboo, Chinese ink, 2006.

The park around the Tiger Hill guides its visitors on paved paths to bridges, ponds, waterfalls, pavillons, a small residence, and, most of all, stones. Many of them tell a story on their own, like the “Rock of sword-probing” which king He Lü cut in half while testing his new weapon. The “pillow rock” once served a buddhist monk for a nap. On the tiger hills peak resides a 1000-year old masonry pagoda.

汁 百 寺 先 汁 沽 升 格 十 侯 十 升 汁 寺 侯 汁



Figure 7: „mǎ“ - horse, Chinese ink, 2006.

汁百吉先汁沽 升格十侯十升汁吉侯汁



Figure 8: „Equus caballus“, Chinese ink, 2006.

汁 吉 吉 充 沽 沽 升 格 十 庚 十 升 沽 吉 庚 汁



Figure 9: „The Brandenburg Gate“, Chinesise ink, 2005.

汁百吉先汁沽 叶格十侯十叶汁吉侯汁

腮
碧
那



Figure 10: „Hofmann S“, Chinese ink, 2005.

汁百吉先汁沽 升格十侯十升汁吉侯汁



Figure 11: "Rhumba", Chinese ink, 2006.

In classic calligraphy the ink is applied from a straightly held brush in a single stroke. The painting has to flow freely from the calligraphers' hand, and pre-drawings or even corrections are therefore excluded. Slowly rubbing the odor-intensive ink from an ink stone sets the calligrapher into a beneficial meditative mood.

汁百吉先汁沽 汁倍十侯十汁汁吉侯汁



Figure 12: „tán yā“, Chinese ink 2006.

Tango can hardly be put in words, for it is full of contradictions.

汁 旨 寺 先 汁 沽 十 倍 十 侯 十 汁 寺 侯 汁



Figure 13: „Arrigo P“, guache, 2005.

A crowded lecture hall in Bielefeld. In the spotlight of the beginning afternoon sun a new stage emerges for the speakers. Each of their movements gets a meaning. Each scientist an actor, each presentation a piece of pantomime. I got my pens with me. Who knows whether the light will be still good after the coffee break?

汁百吉先汁沽 升格十侯十升汁吉侯汁



Figure 14: „Thinking“, Guache 2005.

汁百吉先汁沽 升格十侯十升汁吉侯汁

外
帝



魏
克

Figure 15: „Weinrich D“, Chinese ink, 2006.

The artists of the “Brücke” movement created their famous “15-minute acts” around 1900, as an answer to established forms of drawing and painting in the studio. In the meantime, life is said to have become even faster. Maybe drawing should be pursued rather on a 15-second basis.

汁百吉先汁沽 升格十侯十升汁吉侯汁



Figure 16: „Strackharn J“, Chinese ink, 2006.

汁百吉先汁沽 升格十侯十升汁吉侯汁



Figure 17: „Buba Mara“, guache, 2005.

汁百吉先汁沽 什恰十僕十什汁吉僕汁



Figure 18: „Eckstein R“, guache, 2006.

汁 百 寺 先 汁 沽 升 格 十 侯 十 升 汁 寺 侯 汁

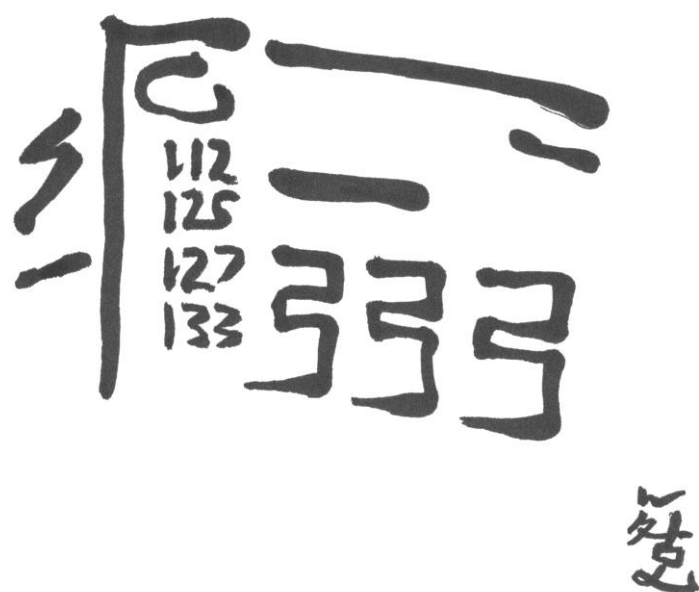


Figure 19: „Esso“, Chinese ink, 2006.



Figure 20: „Imago“, Chinese ink, 2006.

īmāg'ō, n. (pl. -gines, pr. -jīnēz, or -gos). Final & perfect stage of insect e.g. butterfly.

汁百吉先汁沽 叶格十侯十叶汁吉侯汁



Figure 21: „ān há zhàn, S26“, Chinese ink, 2006.

Impressions were caught in midst of life. The often unsuspecting models were moving in their natural habitat. In this volatile moment, a few lines had to suffice to catch everything.

汁百吉先汁沽 升格十侯十升汁吉侯汁

枯
大
学



Figure 22: „Loni“, Chinese ink, 2007.

汁百吉先汁沽 什倍十僕十什汁吉僕汁

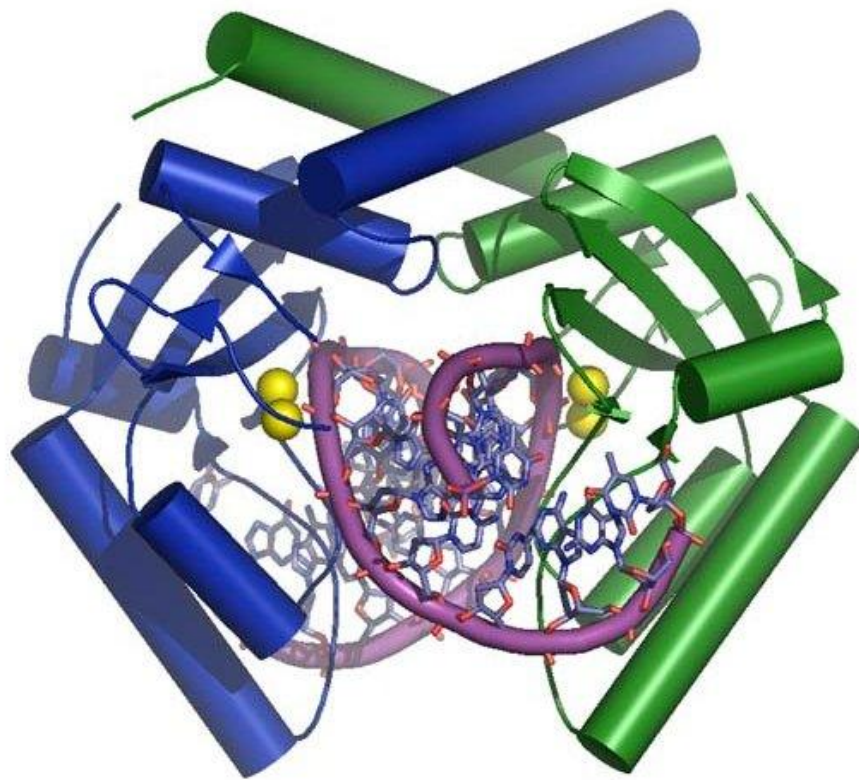


Figure 23: „PvuII Endonuclease“, Stefan Günther, PyMOL, 2004.

汁百吉先汁沽 升格十侯十升汁吉侯汁

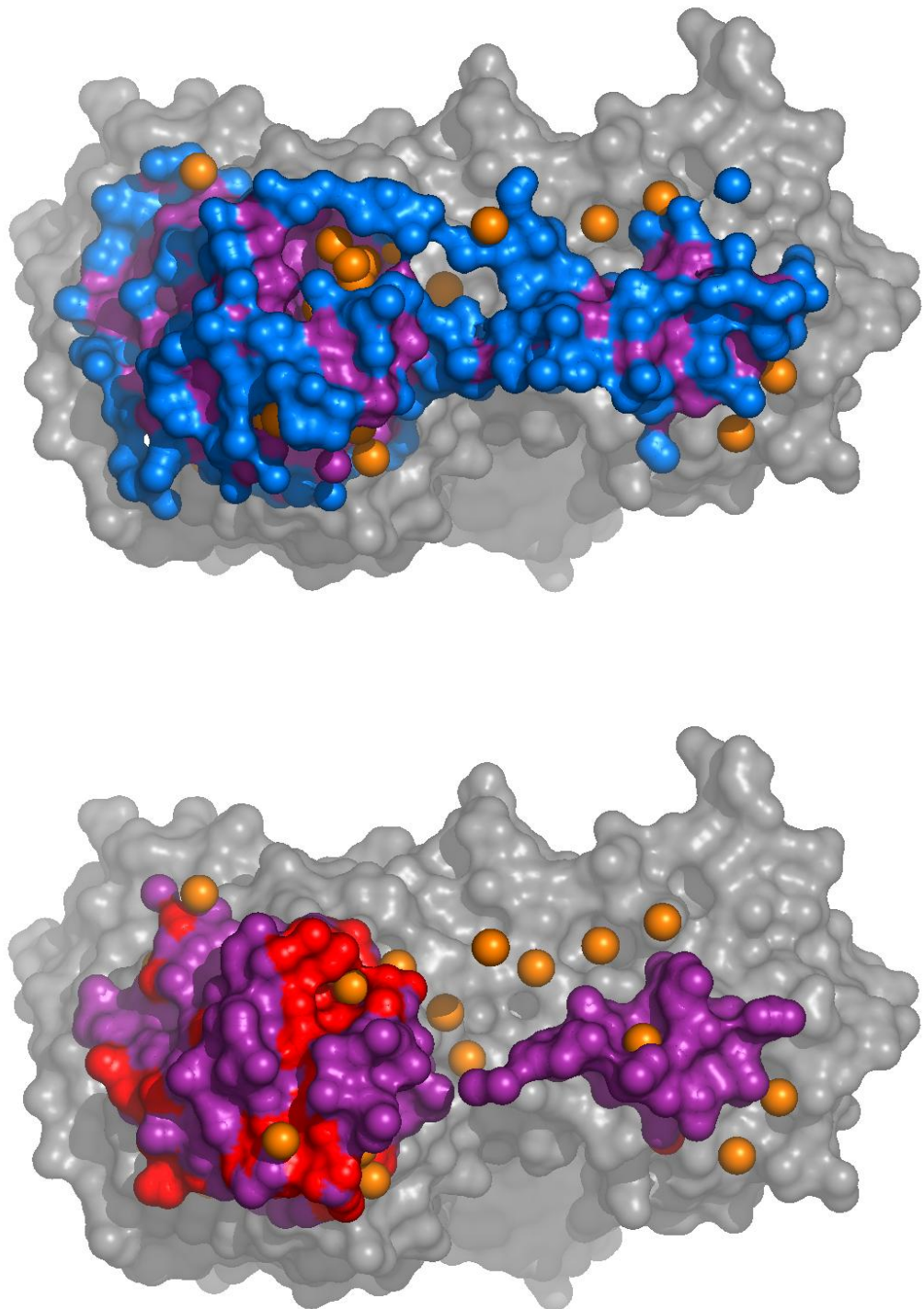


Figure 24: „Packing density of protein structures “, PyMOL, 2003.

汁百吉先汁沽 升格十侯十升汁吉侯汁



Figure 25: „Myoglobin“, guache, 2005.

汁百吉先汁沽 升格十侯十升汁吉侯汁



Figure 26: „Bacteriorhodopsin“, guache, 2005.

汁百吉先汁沽 汁谷十侯十汁沽吉侯汁



Figure 27: „SH3-Domain“, Chinese ink, 2005.

If I had fully mastered the calligraphic technique, could I still pursue science? If I was fully committed to bioinformatics, could I still fuse it with art? I want to strive for what lies beyond these double horizons. Formulate taoist wisdom in the clarity of algorithmic pseudocode. Set alight protein families in the opulence of Zhang Yimou's movies. Paint the Schrödinger orbitals such that an analphabet can grasp their meaning.

Curriculum vitae

Kristian Mikael Rother
born April 11th, 1977 in Berlin
Nationality: Finnish/German
www.rubor.de

Academic Career

10/2006 – 9/2009: Marie Curie Fellowship in the 6th EU Research Training Network; Postdoc in the lab of Dr. habil. Janusz Bujnicki, IIMCB Warszawa

7/2002 - 5/2006: PhD in der Protein Strukturre Theory lab of Prof. Dr. Cornelius Frömmel, Charité Berlin; Thesis: Efficient use of a Protein Structure Annotation Database.

10/1996 - 3/2002: Diploma of Biochemistry at the Free University Berlin; Thesis: Packing density and packing defects of protein structures.

Exhibitions

Visual Reflections on Science, ISMB/ECCB, Stockholm (2009). 1st Prize.

Proteinkalligraphie. Institute of Biochemistry, FU Berlin (2006).

Missing Link - Art Meets Biomedicine, The Rubelle & Norman Schafner Gallery, Brooklyn, N.Y. (2006).

arts & science in vitro - Der seltsame Tanz der sozialen Amöbe. Public act at the Institute of Biochemistry, Charité Berlin (2006).

The Missing Link – Communication in Science and Arts. BMM Charité (2005).