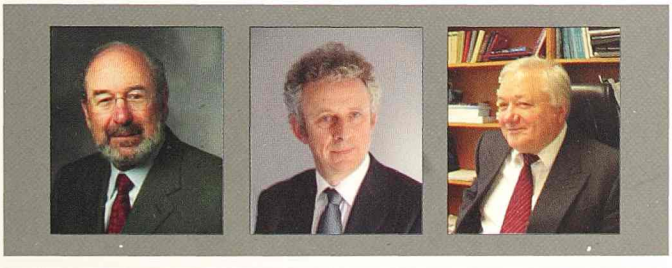
THE KING FAISAL MEMORIAL ARTICLES IN MEDICINE AND SCIENCE IX

جائزة الملك فيصل العالمية
King Faisal International Prize

THE KING FAISAL MEMORIAL ARTICLES IN MEDICINE AND SCIENCE IX



King Abdullah Ibn Abd Al-Aziz
Custodian of the Two Holy Mosques
Supreme Chairman of King Faisal
Foundation

King Abdullah Ibn Abd Al-Aziz
Custodian of the Two Holy Mosques
Supreme Chairman of King Faisal
Foundation

HRH Prince Sultan Ibn Abd Al-Aziz
Crown Prince, Deputy Premier
Minister of Defense and Aviation and
Inspector General

Contents

INTRODUCTION

The King Faisal Foundation continues the traditions of Arabic and Islamic philanthropy, as they were revitalized in modem times by King Faisal. The life and work of the late King Faisal bin Abd Al-Aziz, son of Saudi Arabia's founder and the Kingdom's third monarch, were commemorated by his eight sons through the establishment of the Foundation in 1976, the year following his death. Of the many philanthropic activities of the Foundation, the inception of King Faisal International Prizes for Medicine in 1981 and for Science in 1982 will be of particular interest to the reader of this book. These prizes were modeled on prizes for Service to Islam, Islamic Studies and Arabic Literature which were established in 1977. At present, the Prize in each of the five categories consists of a certificate summarizing the laureate's work that is hand-written in Diwani calligraphy; a commemorative 24-carat, 200 gram gold medal, uniquely cast for each Prize and bearing the likeness of the late King Faisal; and a cash endowment of SR750,000 (US$200,000). Co-winners in any category share the monetary award. The Prizes are awarded during a ceremony in Riyadh, Saudi Arabia, under the auspices of the Custodian of the Two Holy

Mosques, the King of Saudi Arabia.

Nominations for the Prizes are accepted from academic institutions, research centers, professional organizations and other learned circles worldwide, as well as from previous laureatues. After preselection by expert reviewers, the short-listed works are submitted for further, detailed evaluation by carefully selected international referees. Autonomous, international specialist selection committees are then convened at the headquarters of the King Faisal Foundation in Riyadh each year in January to make the final decisions. The selections are based solely on merit, earning the King Faisal International Prize the distinction of being among the most prestigious of international awards to physicians and scientists who have made exceptionally outstanding advances which benefit all of humanity.

(Excerpt from Introduction to 'Articles in Medicine and Science 1"

by H.R.H. Khaled Al Faisal, Chairman of the Prize Board and Director General of King Faisal Foundation)

King Faisal International Prize for Medicine

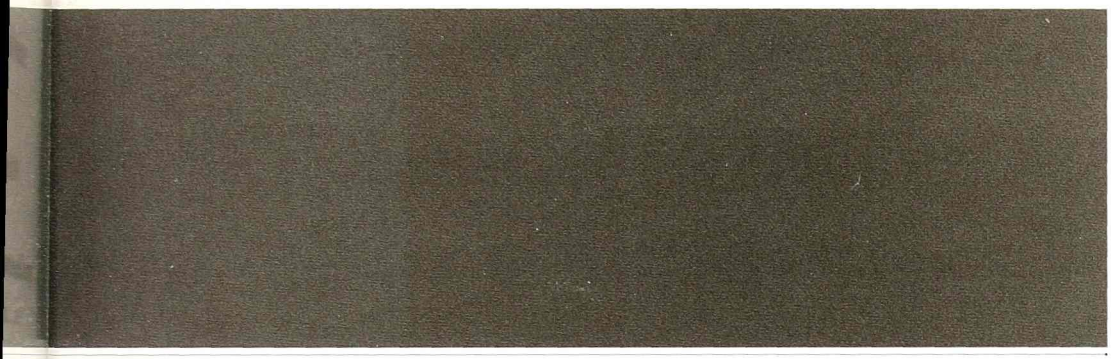
2009 Prize Awards in Medicine and Science

The 2009 awards were presented in March 2009

The Prize for Medicine (Topic: Molecular Targeted Therapy) has been awarded to: Professor Ronald Levy (USA)

Professor Ronald Levy has been recognized for his pioneering studies in Cancer Immunotherapy. Almost 30 years ago he developed antibodies that could distinguish between malignant and benign tumor cells. This created a tool for diagnosis and therapy. He showed in 1980 that monoclonal antibodies against a tumor-associated antigen could be employed to treat patients with B cell lymphoma. Many of these patients responded well to the treatment. Professor Levy generated a humanized monoclonal antibody against a specific protein expressed on malignant B cell. In the last 11 years the drug he discovered has been used to treat a large number of lymphoma patients with remarkable remission and survival results.

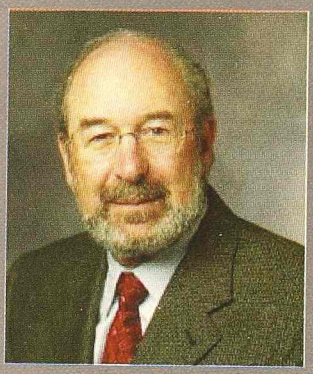
In his more recent work professor Levy has used the patient's immune system to mount immune responses against the patient's own cancer cells. The ability of a patient to elicit immune responses against his/her own tumor cells by vaccinating the patient against the malignancy is likely to become a more effective way for longer lasting and more efficient therapy.



The Prize for Science (Topic: Physics) has been jointly awarded to: Professor Sir Richard H. Friend (UK) and Professor Rashid A. Sunyaev (Russia)

Professor Richard Friend pioneered the physics and engineering of semi-conductor devices made of plastic materials. The fabrication of these devices and circuits by ink jet printing that he has introduced is radically different from the traditional manufacturing technology and has paved the way for the development of a wide range of applications of plastic electronic and photonic devices.

Professor Rashid Sunyaev has made pioneering and fundamental contributions to astrophysics and cosmology. His theoretical work on the cosmic background radiation laid the foundation for the observational exploration of the structure of the universe. His work on black holes and binary stars was critical in advancing the field of x-ray astronomy.



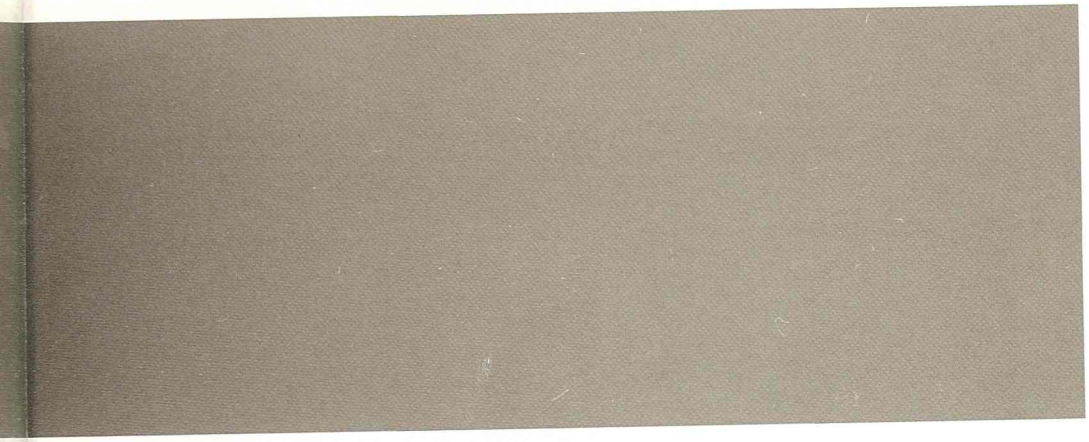
Professor Ronald Levy

Winner of the 2009
King Faisal International Prize for
Medicine

Synopsis of Achievements

Professor Ronald Levy was born on December 6, 1941, in Carmel, California (USA). He received his bachelor's degree (A.B.) from Harvard University (1963) and M.D. from Stanford University (1968) and completed his internship and residency in internal medicine at Massachusetts General Hospital in Boston (1968-1970). He joined the Faculty of the Stanford University School of Medicine in 1975 and is currently the Robert K. and Helen K. Summy Professor, Frank and Else Schilling American Cancer Society Clinical Research Professor and Chief of the Oncology Division in the Department of Medicine at Stanford University.

Professor Levy's exceptional contributions have been instrumental in establishing the role of monoclonal antibodies in the treatment of B-cell lymphomas. Over the past 30 years, he discovered that it was possible to generate antibodies that would recognize specific receptors on these cancer cells and tag them for destruction; he conceived the idea of using these antibodies as molecular targeted drugs to kill the tumor cells and saw his



idea move all the way to an FDA approved drug, Rituxan, now treating half a million people each year with B-cell lymphoma. He has also developed and worked on the idea of a customized vaccine based on the patient's own tumor cells and is using state-of-the art genomic technologies to evaluate genetic signatures that would allow prediction of a response to treatment.

Professor Levy, or the "antibody hero" as he is sometimes nick-named, has received numerous awards and honors, including nearly every major award in cancer research, as well as fellowships and memberships of learned societies and an impressive list of invited lectureships. He has published more than 150 papers in premier medical journals and has been a visiting professor at the Memorial Sloane Kettering Cancer Center in New York and the universities of Texas Southwestern, Minnesota, Nebraska and Miami.

Targeted Therapy for Lymphoma
The Story of the First Monoclonal Antibody for the Treatment of Cancer

Ronald Levy
Department of Medicine
Stanford University School of Medicine
Stanford, California, USA

Thirty-five years ago, when I began my independent research career at Stanford University, there was hope that the immune system could be used to treat cancer. But at that time no one knew how to get the immune system to recognize and reject cancer cells. For instance it was generally believed that antibodies, proteins in the blood produced by the immune system to recognize foreign invaders, would not be powerful enough to attack cancer cells. But subsequent scientific discoveries ultimately changed that view. Today we have antibody treatments for many types of cancer, including lymphoma, leukemia, breast cancer and colon cancer. More are on the way.

Antibodies are one of the ways that we recognize and destroy foreign invaders such as viruses and bacteria. Cancer is not a foreign invader, it is part of our own body, but it is an invader, nevertheless.

My involvement in this story began during medical school when I spent time at the Weizmann Institute of Science, in Israel. There, Dr. Norman Klinman introduced me to a method for cloning the B cells that make antibodies (1). Normally antibodies in the blood are a complex mixture made as part of the immune response to foreign invaders such as viruses and bacteria. By transferring small numbers of the B cells from an immune animal to another normal animal he could isolate small fragments of tissue from the spleen that contained only one of the donor cells. These fragments

Monoclonal Antibodies For Cancer

Year Approved	Product	Target	Original Indication
1997	Rituxan	CD20	B Cell Lymphoma
1998	Herceptin	Her2	Breast Cancer
2000	Mylotarg	CD33	AML
2001	Campath	CD52	CLL
2002	Zevalin	CD20	B Cell Lymphoma
2003	Bexxar	CD20	B Cell Lymphoma
2004	Avastin	VEGF	Colon Cancer
2004	Erbitux	EGFR	Colon Cancer
2006	Vectibix	EGFR	Colon Cancer

could be simulated to make the antibody coming from that one donor cell (monoclonal antibodies).

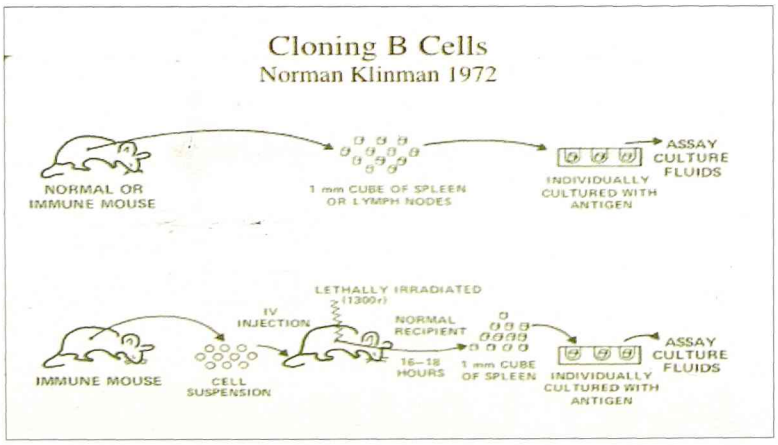


Figure 1. Cloning the cells that make antibodies.

Fragments of spleens can be cultured in vitro and stimulated to make antibodies.

When spleen cells from immune animals are transferred to normal animals the spleen fragments from recipient animals contain only a single donor B cell and make a single (monoclonal) antibody).

Each of these antibody proteins recognized a unique target and each antibody could therefore be used as a tool to discover new targets. Later we used this cell cloning method to make antibodies against human leukemia and lymphoma cells and we were able to define potential new targets for diagnosing and treating these diseases (2). The problem was that the cells that produced the antibodies could not survive very long outside the body and therefore the supply of each antibody was very limited.

Then, in 1975 a revolutionary discovery by Kohler and Milstein, two scientists working together in England, changed all this (3). They invented a way of capturing antibody- producing cells by fusing them to myeloma cells, a type of cancer cells that had the machinery for making antibody proteins, but also had the property of all cancer cells - they could live forever. These hybrid cells - called hybidomas, could make the antibody product of the immune cell in quantities that were unlimited. Suddenly we realized that antibodies could be eventually tested as a treatment for cancer. Kohler and Milstein received the Nobel Prize for this work in 1984.

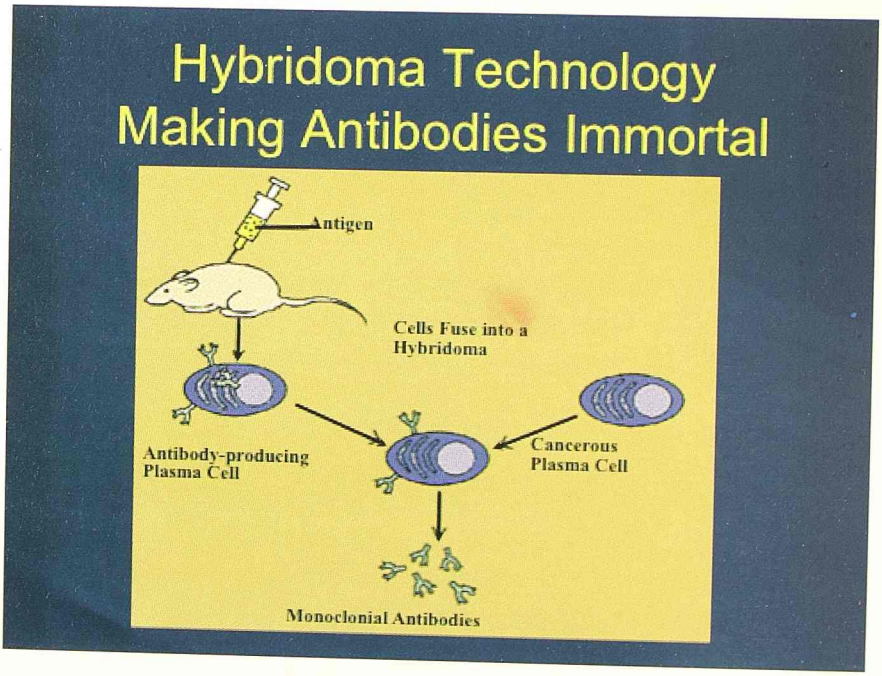


Figure 2. Monoclonal Antibody Production by Hybridoma Technology. Spleen cells from immune animals were fused to myeloma tumor cells. The resulting hybrid cells (hybridomas) make the specific antibody and live forever.

A Perfect Target

We set out to find targets on cancer cells that antibodies could recognize. Our first target was one that we knew to be present on lymphoma tumor cells and not on normal cells of the body. Cancer starts from one original cell that loses the ability to control its growth. Lymphoma is the cancer of lymphocytes and these cells each make a different antibody protein. The part of this protein that is different from one to the next is called the "idiotype".

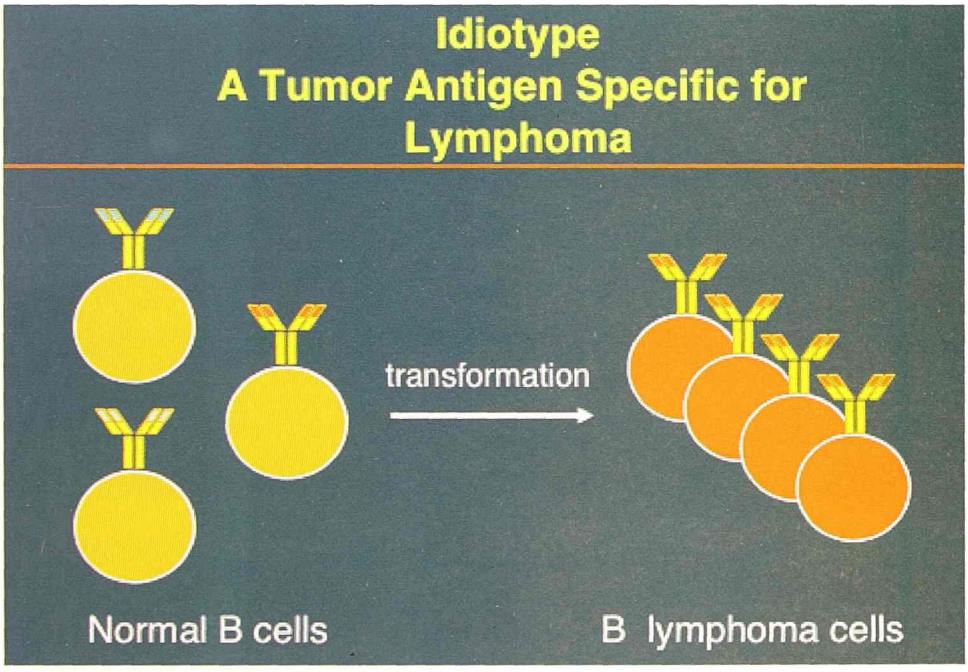


Figure 3. Idiotype, A Specific Target on Lymphoma Tumor Cells.
Each normal B cell makes only one antibody. Lymphoma arises from one of these normal B cells and all the tumor cells make the same exact antibody. This antibody (idiotype) can be a marker of the tumor and a target for therapy.

The concept that we worked on was to make monoclonal antibodies that could recognize the unique idiotype present on each patient's tumor- antibodies against antibodies.

We made antibodies against this target, tested them on mouse lymphoma tumors, and it worked. The mice could be cured by this

treatment. By then, researchers in other parts of the world, including George and Freda Stevenson, a husband and wife team in England were getting similar results (4).

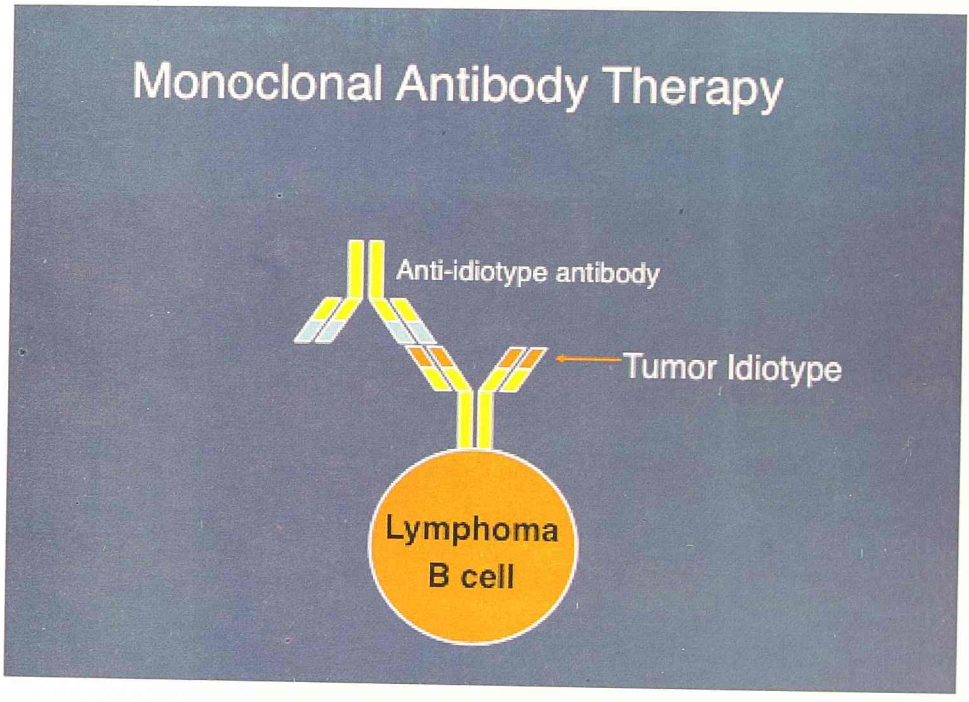


Figure 4. Anti-Idioytpe Therapy.
Antibodies can be made that recognize the unique idiotype present on each tumor. Such antibodies can be used to treat the tumor. A different antibody needs to be made for each patient.

In 1981 Richard Miller, David Maloney and I treated the first lymphoma patient with a monoclonal anti-idiotype antibody and we were amazed by the result. Tumors throughout the patient's body melted away and disappeared completely (5).

Emboldened by this success, we were anxious to repeat it in other patients. However, the task was not simple. The lymphoma tumor of each patient has a different idiotype, and because of this we had to make a custom antibody for every patient. We did this for fifty patients over a period of ten years. Most of these patients had shrinkage of their tumors

and some, including our first patient, had benefit that lasted for a decade or more (6).

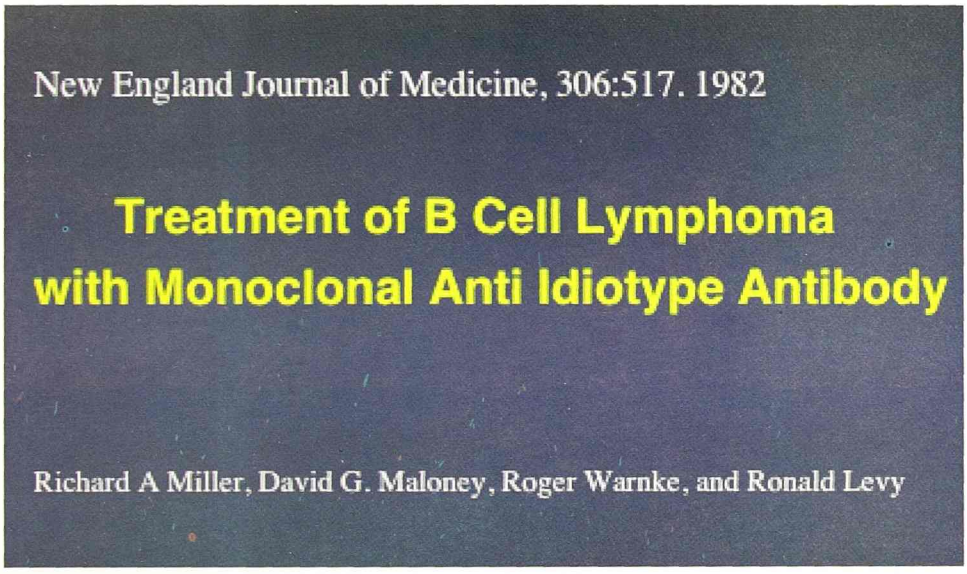


Figure 5. The First Report of Successful Treatment of Cancer with a Monoclonal Antibody.

Development of Rituximab: A Generic Approach

The customized anti-idiotype approach was not practical for use in a larger numbers of patients. Instead, scientists at Idec Pharmaceutical Company made an antibody against CD20, a target found on all B cell lymphomas. This new approach had the advantage that the same monoclonal antibody might be used to treat all patients with B cell lymphoma.

However, CD20 is also present on normal B lymphocytes, the very cells that make the antibodies that defend us against infections. Therefore, no one knew if such a treatment would be safe. Our first clinical trial showed that Rituximab, the antibody against CD20, worked to treat patients with lymphoma (7-10). Of the several different doses we explored, none was clearly superior, but we chose the highest dose for further development. Our initial results were confirmed and extended

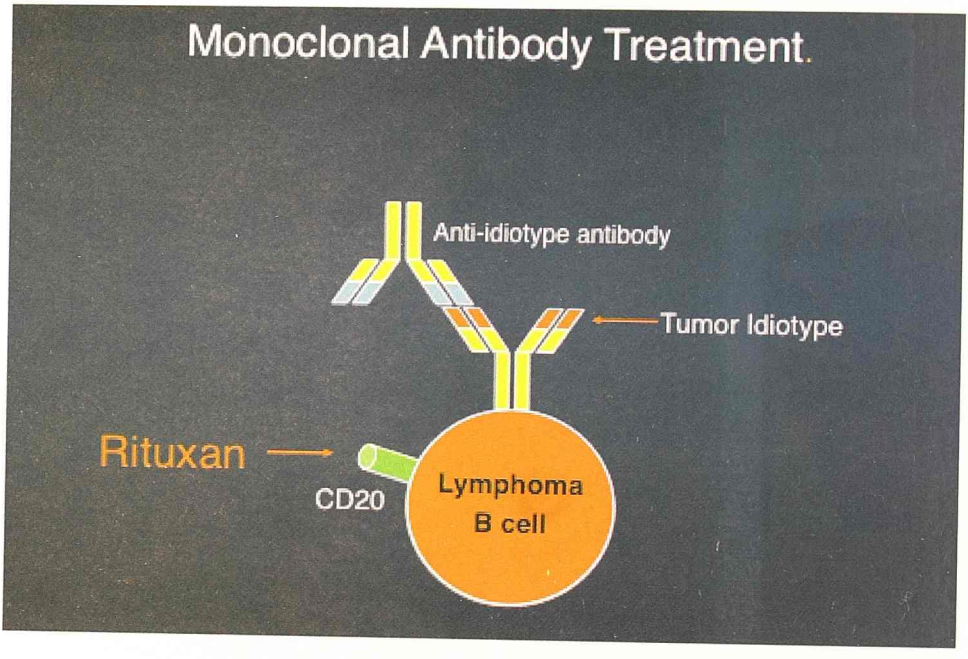
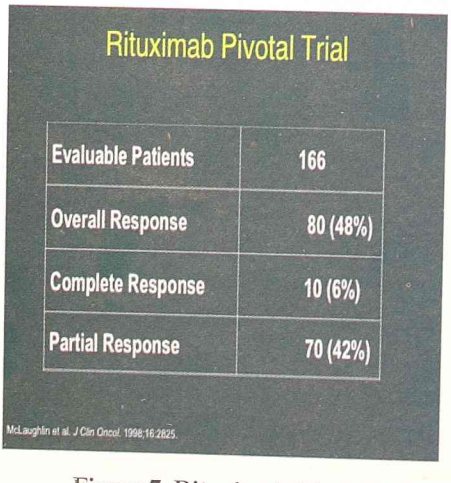


Figure 6. Targeting B cells.
Rituximab, otherwise known as Rituxan, recognizes CD20, a target present on all malignant and normal B lymphocytes.

by investigators throughout our country (11), ultimately leading to an approval of the drug by the U.S. FDA in 1997.



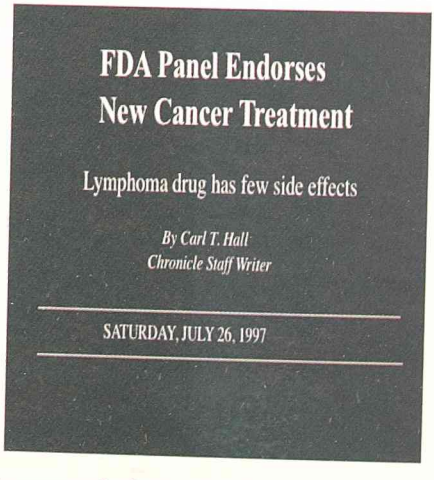


Figure 7. Rituximab Pivotal Trial and FDA approval of Rituximab in 1997.

Despite the presence of CD20, the target of Rituximab, on normal B cells and the fact that these normal cells are depleted from the blood during treatment with Rituximab, the patients do not become immune suppressed and do not suffer an increased rate of infections.

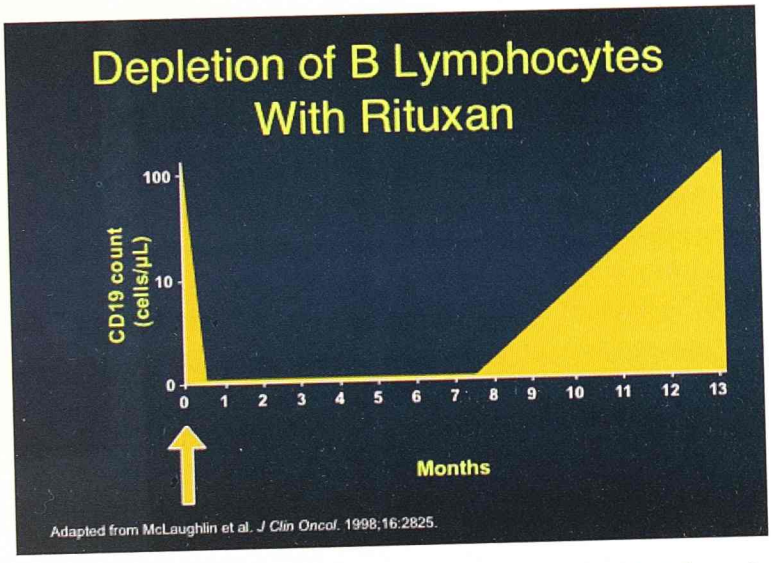


Figure 8. Schematic Illustrating the Disappearance of Normal B Lymphocytes from the Blood of Patients Treated with Rituximab.

This depletion typically lasts for at least six months after a single treatment.

Widespread Use of Rituximab in Lymphoma.

Once the drug was approved and became available, clinical investigators throughout the world began testing it in combination with other lymphoma treatments and in clinical situations that went beyond the narrow initial approval. For instance, investigators in France showed that patients with a more aggressive form of B cell lymphoma responded better and lived longer when Rituximab was added to the standard chemotherapy treatments (12). Other investigators in the U.S. showed that Rituximab could be used as the first treatment for the slow growing type of lymphoma, thus delaying or avoiding the need for chemotherapy (13). Clinical investigators in many different countries have now shown that Rituximab can be used either before, during or after chemotherapy, and in each situation patients are benefited (14-16). Today, Rituximab is used for all patients with B cell lymphoma and it is prolonging the lives of many of them.

There are many questions still to be answered about Rituximab treatment for lymphoma. How do antibodies work? Why are some tumors resistant to antibody treatment? Why do tumors become resistant to antibody treatment? Can new monoclonal antibodies be found against CD20 or against other targets on lymphoma cells to improve on the treatment?

Extended Uses of Rituximab

A recent and very surprising discovery is that Rituximab works to treat auto-immune diseases, non-malignant diseases caused by an over active immune system (17). Presumably it does this by interfering with the B cells that cause these diseases. This list of Rituximab treatable diseases is still growing, but it already includes rheumatoid arthritis, multiple sclerosis, system lupus erythematosis, auto-immune hemolytic anemia, auto-immune thrombocytopenic purpura, and pymphigus vulgaris.

A Model For the Future

This story shows that new discoveries in the laboratory can lead to new treatments in the clinic. It depended on a combination of scientific intuition and hard work by many different individuals and organizations and a good deal of luck. Granting agencies fund the early work. Scientists make the discoveries. Investors and companies take the financial risk. Clinical investigators test the new treatments in patients. Government regulatory agencies assure that the treatments are safe and effective. Doctors adopt the treatments and use them. New uses for the treatments are then found.

Many different people working in different parts of the world shared their information with each other and contributed to the development of monoclonal antibody therapy for cancer. Cancer remains a major problem for all of us. New discoveries need to be made, new preventative measures need to be found, and new treatments need to be developed. Hopefully monoclonal antibodies will be a part of that ultimate solution.

References

1. Klinman, N. R., and J. L. Press. 1975. The characterization fo the B-cell repertoire specific for the 2,4-dinitrophenyl and 2,4,6-trinitrophenyl determinants in neonatal BALB/c mice. J Exp Med 141:1133.
2. Levy, R., and J. Dilley. 1977. The in vitro antibody response to cell surface antigens. II. Monoclonal antibodies to human leukemia cells. J Immunol 119:394.
3. Kohler, G., and C. Milstein. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 256:495.
4. Hamblin, T. J., A. K. Abdul-Ahad, J. Gordon, F. K. Stevenson, and G. T. Stevenson. 1980. Preliminary experience in treating lymphocytic leukaemia with antibody to immunoglobulin idiotypes on the cell surfaces. Br J Cancer 42:495.
5. Miller, R. A., D. G. Maloney, R. Warnke, and R. Levy. 1982. Treatment of B-cell lymphoma with monoclonal anti-idiotype antibody. N Engl J Med 306:517.
6. Davis, T. A., D. G. Maloney, D. K. Czerwinski, T. M. Liles, and R. Levy. 1998. Anti-idiotype antibodies can induce long-term complete remissions in non-Hodgkin's lymphoma without eradicating the malignant clone. Blood 92:1184.
7. Maloney, D. G., A. J. Grillo-Lopez, D. J. Bodkin, C. A. White, T. M. Liles, I. Royston, C. Varns, J. Rosenberg, and R. Levy. 1997. IDEC-C2B8: results of a phase I multiple-dose trial in patients with relapsed non-Hodgkin's lymphoma. J Clin Oncol 15:3266.
8. Maloney, D. G., A. J. Grillo-Lopez, C. A. White, D. Bodkin, R. J. Schilder, J. A. Neidhart, N. Janakiraman, K. A. Foon, T. M. Liles, B. K. Dallaire, K. Wey, I. Royston, T. Davis, and R. Levy. 1997. IDEC-C2B8 (Rituximab) anti-CD20 monoclonal antibody therapy in patients with relapsed low-grade non-Hodgkin's lymphoma. Blood 90:2188.
9. Davis, T. A., C. A. White, A. J. Grillo-Lopez, W. S. Velasquez, B. Link, D. G. Maloney, R. O. Dillman, M. E. Williams, A. Mohrbacher, R. Weaver, S. Dowden, and R. Levy. 1999. Single-agent monoclonal antibody efficacy in bulky non-Hodgkin's lymphoma: results of a phase II trial of rituximab. J Clin Oncol 17:1851.
10. Davis, T. A., A. J. Grillo-Lopez, C. A. White, P. McLaughlin, M. S.

Czuczman, B. K. Link, D. G. Maloney, R. L. Weaver, J. Rosenberg, and R. Levy. 2000. Rituximab anti-CD20 monoclonal antibody therapy in non-Hodgkin's lymphoma: safety and efficacy of re-treatment. J Clin Oncol 18:3135.

11. McLaughlin, P., A. J. Grillo-Lopez, B. K. Link, R. Levy, M. S. Czuczman, M. E. Williams, M. R. Heyman, I. Bence-Bruckler, C. A. White, F. Cabanillas, V. Jain, A. D. Ho, J. Lister, K. Wey, D. Shen, and B. K. Dallaire. 1998. Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program. J Clin Oncol 16:2825.
12. Coiffier, B., E. Lepage, J. Briere, R. Herbrecht, H. Tilly, R. Bouabdallah, P. Morel, E. Van Den Neste, G. Salles, P. Gaulard, F. Reyes, P. Lederlin, and C. Gisselbrecht. 2002. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. N Engl J Med 346:235.
13. Hainsworth, J. D., H. A. Burris, 3rd, L. H. Morrissey, S. Litchy, D. C. Scullin, Jr., J. D. Bearden, 3rd, P. Richards, and F. A. Greco. 2000. Rituximab monoclonal antibody as initial systemic therapy for patients with low-grade non-Hodgkin lymphoma. Blood 95:3052.
14. Habermann, T. M., E. A. Weller, V. A. Morrison, R. D. Gascoyne, P. A. Cassileth, J. B. Cohn, S. R. Dakhil, B. Woda, R. I. Fisher, B. A. Peterson, and S. J. Horning. 2006. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. J Clin Oncol 24:3121.
15. Herold, M., A. Haas, S. Srock, S. Neser, K. H. Al-Ali, A. Neubauer, G. Dolken, R. Naumann, W. Knauf, M. Freund, R. Rohrberg, K. Hoffken, A. Franke, T. Ittel, E. Kettner, U. Haak, U. Mey, C. Klinkenstein, M. Assmann, and U. von Grunhagen. 2007. Rituximab added to first-line mitoxantrone, chlorambucil, and prednisolone chemotherapy followed by interferon maintenance prolongs survival in patients with advanced follicular lymphoma: an East German Study Group Hematology and Oncology Study. J Clin Oncol 25:1986.
16. Marcus, R., K. Imrie, P. Solal-Celigny, J. V. Catalano, A. Dmoszynska, J. C. Raposo, F. C. Offner, J. Gomez-Codina, A. Belch,

D. Cunningham, E. Wassner-Fritsch, and G. Stein. 2008. Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. J Clin Oncol 26:4579.

17. Edwards, J. C., L. Szczepanski, J. Szechinski, A. Filipowicz-Sosnowska, P. Emery, D. R. Close, R. M. Stevens, and T. Shaw. 2004. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med 350:2572.

King Faisal International Prize

for Science

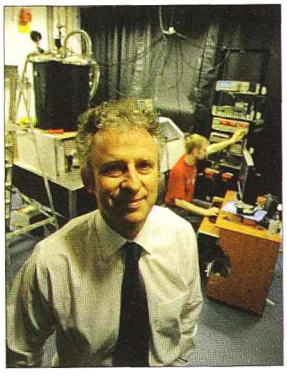
Professor Sir Richard H. Friend

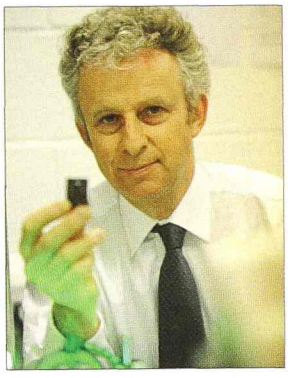
Co-Winner of the 2009 King Faisal International Prize for Science

Synopsis of Achievements

Professor Sir Richard Friend is one of today's great physicists. Born on January 18, 1953 in London, he graduated from Trinity College in 1974 with a BA (First Class) in Theoretical Physics and obtained his Ph.D. from the Cavendish Laboratory at the University of Cambridge in 1978. He has been on the Faculty of the Department of Physics at Cambridge since 1980 and is currently holding the prestigious Cavendish Professorship of Physics at that university. He is also a Fellow of St. John's College and Chairman of the Council of the School of Physical Sciences at the University of Cambridge and the Tan Chin Tuan Centennial Professor at the National University of Singapore. Moreover, he is the founder and Chief Scientist of Cambridge Display Technology Ltd. and Consultant at Plastic Logic Ltd.

Professor Friend's pioneering work on the semiconductor physics of conjugated polymers has had a profound impact on physics and beyond. He has essentially invented a new type of electronics using organic





semiconductors and has persisted with their development into polymer light-emitting diodes that are now widely used and offer the potential of cheaper, larger and flexible displays. He continues to develop polymer photovoltaics and directly printed polymer transistors. He has authored around 700 publications in peer-reviewed journals, and more than 20 patents. The Institute for Scientific Information has identified him as the most cited physicist in the UK for the decade 1990-1999 and is currently one of the two most cited physicists in his country (~ 39,000 citations).

Professor Friend is the recipient of numerous awards and honors. He is a Fellow of the Royal Society of London, Fellow of the Royal Society of Engineering, Honorary Fellow of the Royal Society of Chemistry, Honorary Fellow of the Institute of Physics, Honorary Fellow of Trinity College (Cambridge) and Honorary Fellow of the University of Wales (Bangor). He holds honorary doctorate degrees from the universities of Linkoping (Sweden), Mons-Hainaut (Belgium) Heriot-Watt (Edinburgh) and Radboud Nijmegan (Netherlands). He is also the recipient of many prizes including the prestigious Rumford Medal of the Royal Society of London, Silver Medal of the Royal Academy of Engineering, Faraday Medal of the Institute of Electrical Engineers, Gold Medal of the European Material Research Society and Descartes Prize of the European Commission. He was knighted in 2003 for his services to physics.

The Science and Technology of Plastic Electronics

Sir Richard H. Friend
Cavendish Laboratory
J.J. Thomson Avenue
Cambridge, CB3 OHE, U.K.

Introduction: Biology uses carbon-based molecules to construct elaborate and functional structures, and these are assembled using specific cooperative intermolecular interactions. In contrast, semiconductor technology is currently based around inorganic semiconductors, principally silicon. High temperature processing methods are generally used, requiring that devices be constructed within single crystals of semiconductor, and thus be restricted to small active areas. Though this technology delivers remarkable performance, the restrictions on large-area processing are well-recognised. Large area thin film silicon transistor technology has been built around the use of amorphous silicon-hydrogen alloys that can be deposited by decomposition of silane at temperatures compatible with deposition onto glass substrates. However, there are few inorganic semiconductors that show useful electronic properties when deposited as amorphous structures.

Over the past two decades we and others have discovered that molecular materials can show useful semiconductor properties and can also be incorporated into practical manufacturing processes. They now show real promise where large areas of active semiconductor devices are required, principally in displays, where they can be used both as light-emitting diodes and also as the transistors required to construct the active-matrix addressing to each of the pixels in the display, and now also as the

active semiconductors in photovoltaic diodes, solar cells.

Semiconductor devices generally need the semiconductor to be processed to form uniform thickness thin films, and this is now successfully accomplished in a variety of ways. Vacuum sublimation onto a cold substrate can allow deposition of amorphous films, that, with appropriate choice of molecular structure, can be sufficiently stable against recrystallisation. The pioneering work that set up this approach was due to Tang and co-workers at Kodak [1] who demonstrated that multiple layers of molecular semiconductors could be deposited, and that the heterojunctions between these layers could be used to control device operation, as discussed below. The second approach, as used in our laboratory, is to use materials that are film-forming from solution in common solvents. Polymers, or composites containing polymers are generally required to achieve this, and the development of this class of materials has enabled this solution-processing route, starting from our demonstration of field-effect transistors devices [2] and our finding of light-emission in polymer semiconductor diodes [3].

The electronic properties of organic molecules with semiconducting properties were first investigated in the 1960s with studies then principally of single crystals of molecules such as anthracene [4], with first measurements of electroluminescence from these crystals reported [5,6]. The ability of these molecular semiconductors to allow electronic transport, and to show charge photogeneration in appropriate structures was picked up and developed for use in electrophotography, initially as a replacement for the photoconductive selenium first used Xerox. It was found possible to incorporate molecular semiconductors that were not directly film-forming within polymer hosts and in this way process the photoconductive layers onto the drum used for image writing and subsequent ink capture and transfer.The vast majority of laser printers and copiers now use these materials systems.

The current interest in semiconducting organic materials draws both on the developments of the molecular semiconductors used in electrophotography and also from the field of 'conducting' π-conjugated polymers that started with the discovery of metallic conduction in charge-transfer complexes formed between polyacetylene and oxidising agents such as iodine [7]. The use of organic semiconductors, both molecular and polymeric, as the active materials within conventional semiconductor device structures forms the material for this article.

Organic Semiconductor materials

Carbon-based molecular semiconductors have a common structural feature – that the carbon-carbon bonding within the molecule or along the polymer backbone can be described as the alternation of carbon-carbon single bonds and carbon-carbon double bonds. This feature is termed 'conjugation' and such materials are commonly described as conjugated polymers or molecules. Carbon-carbon double bonds are formed when carbon is bonded through sp^2 hybrid orbitals which result to three covalent σ-bonds within a plane, leaving the out-of-plane p_z orbital non-hybridised. Figure 1 shows the structures of some semiconducting polymers.

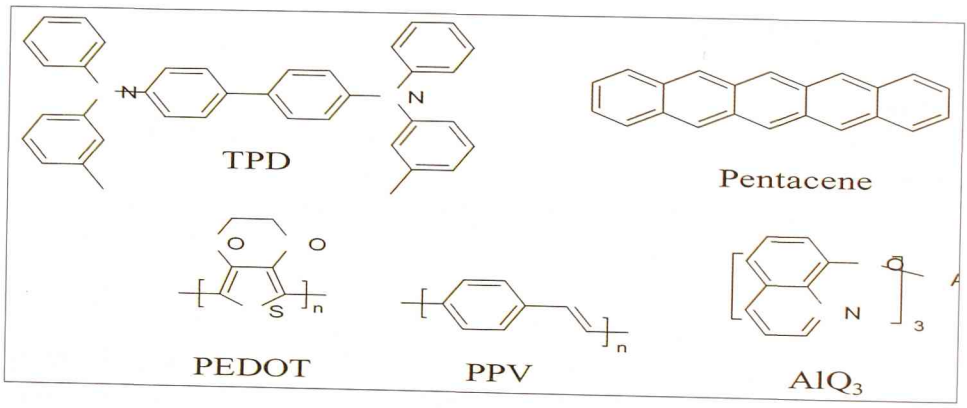


Figure 1 Structures of some polymeric semiconductors. PEDOT is a derivative of polythiophene, is relatively stable when oxidatively doped and is used as a charge-transfer complex with polystyrene sulphonic acid, PSS, as a (poor) metallic conductor; PPV is an example of a luminescent conjugated polymer that can be used as the emissive material in polymer LEDS.

Overlap of this pz orbital with pz orbitals on adjacent carbons allows formation of π-molecular orbitals which are substantially delocalised over the molecule and along segments of the polymer chain. The filled π-bonding orbitals form the valence states, and the empty π*-antibonding orbitals form the conduction states. The formation of well-delocalised conduction and valence band states is one of several necessary conditions to obtain useful semiconducting behaviour. It is important to note that the σ 'single' bonds formed by the sp2 hybrid

orbitals also play an important role: because these are much stronger bonds than those formed with the π electrons they generally keep intact the molecular structure in the presence of excitations of the π electrons.

As mentioned above, it is common to prepare organic semiconductors with disordered structures, either as rapidly-cooled vacuum-sublimed films or as solution-processed films, and this lack of structural regularity must clearly affect electronic properties. However, inter-molecular disorder and intra-chain disorder due to chain twists and torsions does not generally cause changes to the covalent intra-molecular bonding, and therefore does not create 'defect' electronic states that are associated with dangling bonds or dislocation lines in inorganic semiconductors. Thus, 'nonbonding' states within the semiconductor gap (that plague the performance on inorganic semiconductors) are not generally formed. This makes possible very clean interfaces, for example between dissimilar organic semiconductors. Disorder does however affect strongly the transport properties of organic semiconductors, and both charged and neutral excitations can be localised due to disorder. This is a particular problem for organic semiconductors used in FETs where carrier mobility is important.

Processing

The general requirement for organic semiconductor processing is the production of a thin film of uniform thickness. A secondary requirement is for in-plane patterning, as required for example in the production of red, green and blue subpixels. Much of the interest in these materials has been in finding new ways to process them to form useful structures for semiconductor devices, and several very promising methods have been developed. Processing from the liquid phase, usually from a solution in a non-polar solvent, is particularly appealing, because this is scalable to large areas, and can also allow direct patterning by printing.

Polymers have been particularly developed because they are conveniently processed from solution to form uniform thickness thin films. Solution processing is attractive for several reasons. Firstly, it allows elements of 'self-assembly' of one material with respect to another, so that for example, a polymer layer can align onto a pre-ordered substrate (e.g. via a liquid crystalline mesophase), as is

desirable for example in FETs. It can also allow complex morphologies with non-planar interfaces between different semiconductors, and one example of this is the use of controlled de-mixing of two polymers deposited from common solution to form an interpenetrating network of the two polymers. This is needed particularly for PV diodes, as discussed below, where light needs to be absorbed throughout the thickness of the absorbing semiconductor layer but also be absorbed close to the heterojunction between electron- and hole-accepting semiconductors so that charge separation can occur.

Most of the potentially useful applications for organic semiconductor devices also require in-plane patterning. This is the case for full-colour LED displays where separate red, green and blue sub-pixels need to be defined, but is also the case for transistor circuits where each transistor needs to be formed from a single area of organic semiconductor in order to avoid cross-talk between adjacent devices.

There has therefore been great interest in the use of direct printing of semiconducting or doped conducting polymers and this can be realised by their formulation as 'inks' (selection of viscosity and surface tension), so as to be compatible with convenient printing methods. Ink-jet printing has been particularly popular. However, ink-jetted drops as produced by current generation ink-jet print heads have volumes of a few picolitres, and diameters of more than 10 μm. This is adequate for patterning of sub-pixels for LEDs, but not generally useful for defining structures such as FET channels [8]. However, by pre-patterning the surface free energy of the substrate (using techniques such as micro-contact printing) it is possible to steer the flow of the fluid droplet deposited by ink-jet printing, so that its eventual position is very accurately determined.

A further approach that allows the patterning of very narrow channels between source and drain of an FET is to arrange that that the first electrode (printed as a line of droplets from the polymeric metal PEDOT-PSS in aqueous solution) is printed, dried and then rendered hydrophobic (by for example treatment in a CF4 plasma) and the second electrode then printed nearby. As the printed drops spread and reach the first structure they are prevented from wetting and a well-defined channel is set up between the two conducting lines, as shown in figure 2. The channel length between these two electrodes can be arranged to be below 100 nm [9].

Luminescence and excitons

Many molecular semiconductors are highly luminescent, particularly when in dilute solution. The negatively-charged electron and the positively-charged counterbalancing 'hole' generated by photon absorption in the semiconductor are unlikely to separate because molecular semiconductors have low dielectric constants, typically around 3, so that screening of the attractive electrostatic interaction between electron and hole is weak, particularly when compared with inorganic semiconductors such as silicon. In fact, this electrostatic interaction keeps electron and hole confined to not much more than one molecular unit or repeat unit for the molecules and polymers shown in figure 1, and is generally referred to as a neutral excited electronic state or exciton. This is evident from the optical absorption and luminescence shown for the polymer PPV in figure 3, where the electronic transition near 2.5 eV is strongly coupled to vibrational excitations of the polymer chain.

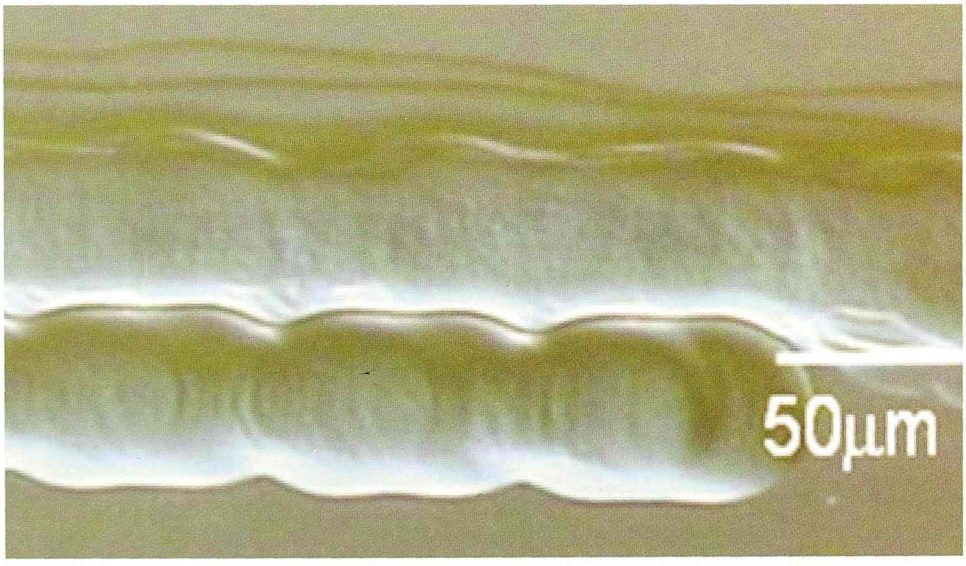


Figure 2 :Optical micrograph of two ink-jet printed lines of the conducting polymer composition PEDOT-PSS. The first line to be printed is surface-treated to repel the second line as it is printed, and the resulting structure can be used as the source and drain of an FET. The effective channel length between the two conducting lines for this structure was found to be below 100 nm [9].

Light-Emitting Diodes

We discovered that very simple diode structures made with a solution-processed layer of semiconducting polymer sandwiched between two electrodes could function as efficient light-emitting diodes [3]. Subsequent development has improved durability and colour control to bring these devices to consumer products as light-emitting pixelated displays. Diode structures can also be manufactured by by vacuum sublimation of molecular semiconductors [1] as pioneered by Tang and colleagues at Kodak, New York.

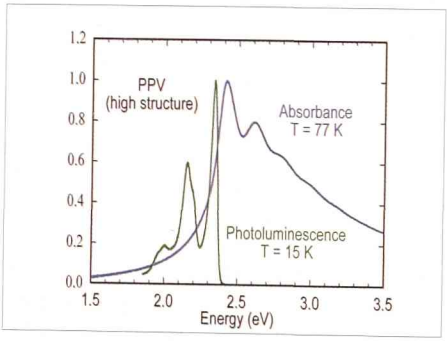


Figure 3a: Optical absorption and luminescence from a well-ordered film of PPV, showing green luminescence

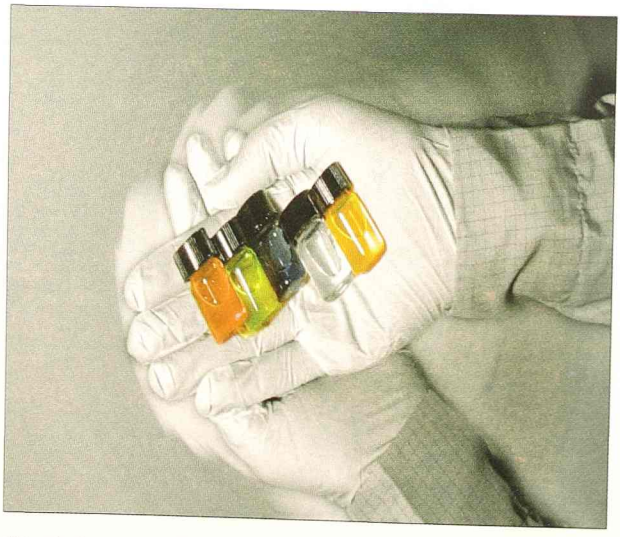
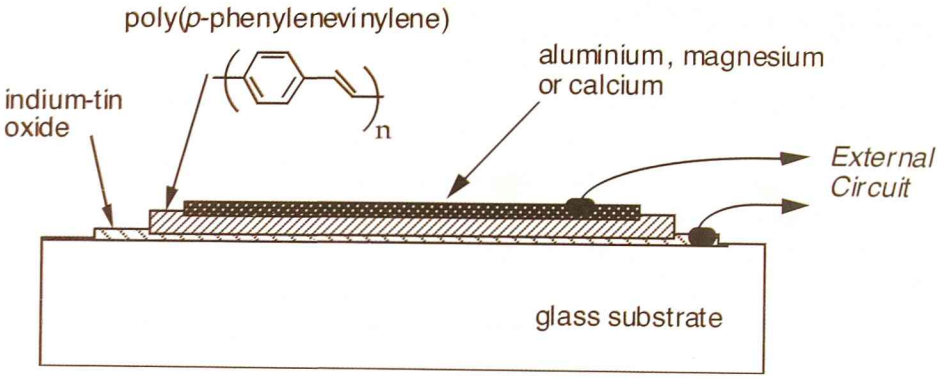


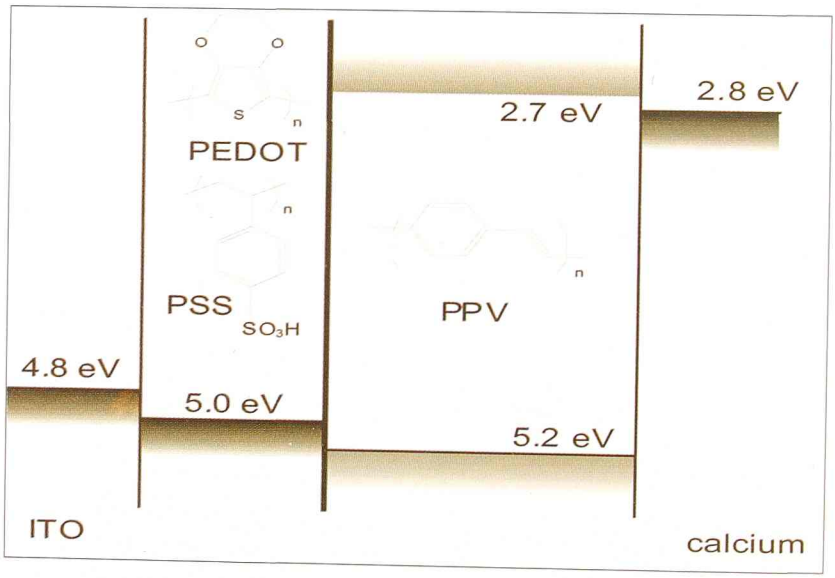
Figure 3b: Solutions of a range of semiconducting polymers with different optical bandgaps and luminescence colours.

These LEDs operate through electron and hole injection from opposite electrodes, electron-hole capture within the bulk of the semiconductor film to form bound excitons, and subsequent radiative emission from the exciton. That these devices work at all, let alone very efficiently, has come as a surprise. The organic semiconductor layers are generally very thin (total thickness of around 100 nm), and interfaces between these and the two inorganic electrodes might be expected to be very troublesome. Indium-tin oxide on a transparent support substrate (usually glass, but also polymer substrates) is generally used as the bottom, hole-injecting electrode. Although it has a high work function, as required for hole injection into the valence band states of the organic semiconductor, it is often coated with a more predictable hole-injecting layer, such as the conducting polymer formulation PEDOT:PSS which is conveniently water soluble. Low work function metals are required for electron injection, and metals such as calcium or magnesium (often alloyed) are widely used. The scheme of energy levels for this model of operation is shown schematically in figure 4(b), in which we have not attempted to modify electrode – semiconductor interfaces to take account of the interfacial interactions. This crude model is moderately successful because the interfacial chemistry at both electrodes is relatively benign, in that it does not introduce 'broken bonds' with energy states deep in the semiconductor gap, but in detail, this has proved to be a very rich area of investigation of both chemical reactions and charge polarisation at interfaces.

The architecture of these LEDs has been considerably developed, particularly those made by direct vacuum sublimation of molecular species. Through control of charge injection and recombination and through the use of phosphorescent organometallic guest emitter species, very high efficiencies are now reported, particularly from industrial groups. For example, Novaled have claimed power efficiency for a green OLED greater than 120 lm/W at 1,000 cd/m2 [10].



(a)



(b)

Figure 4 Structure of simple polymer LED devices: (a) the LED is formed by spin-coating of a solution of the semiconducting polymer, here shown as PPV, onto the anode formed as a layer of transparent indium-tin oxide. The top electrode is then formed by thermal evaporation of a low work-function metal such as calcium onto the semiconductor layer. (b) energy level scheme for a diode formed with a layer of the conducting polymer composition PEDOT-PSS spin-coated onto the indium-tin oxide layer. This acts to improve injection of holes from the ITO into the π valence band of the semiconducting polymer. Electrons are injected from the low workfunction electrode, shown here as calcium, into the π^* conduction band states.

LED Applications

The account above of the development of LED architectures to deliver efficient light emission has been matched by a very substantial industrially-led activity to improve device durability. Both for polymer and for vacuum-sublimed devices lifetimes have been advanced very considerably, with projected lifetimes under drive conditions to half initial brightness (with a starting value usually taken at around 1000 cm/ m2 – about 5 x display brightness) in excess of 100,000 hours. These results are only achieved if there is extremely good encapsulation that prevents ingress of water and oxygen, and at present, requires the use of glass as substrate since it has very much better barrier properties than polymer substrates.

Much of the efforts with organic LEDs have been directed to the manufacture of pixellated full colour displays. These generally require an active matrix of thin-film transistors to drive the correct current through each LED, but the circuit demands are more challenging than those set by liquid crystal displays, because the LED is a current-driven device. The use of laser-annealed polycrystalline silicon transistor arrays has therefore been developed, with the local transistor circuit comprising at the minimum, two transistors per sub-pixel. There are at present many demonstrator displays seen at trade shows, made both with polymers (patterned by ink-jet printing) and, more commonly, vacuum-sublimed molecules. Patterning of the latter however remains very difficult to achieve, and one approach that is currently adopted is to deposit white-emitting diodes, and to use a red/green/blue colour filter on top of these. The disadvantage of this approach is that much of the white light generated is discarded, and efficiency is reduced by a factor of 3 or so. As of early 2008, there is now one commercially-available display on the market, produced by Sony, using vacuum-sublimed molecular films.

Photovoltaic Diodes

There is renewed interest in the scope for using thin-film organic semiconductors in PV diodes. These are potentially very attractive because there is scope for low cost manufacture onto plastic substrates. The basic challenge that has to be met is that optical absorption in such materials generates a bound exciton that will not readily dissociate in the bulk to electron and hole because the coulombic binding energy is substantial (of order 0.5 eV). The approach that is universally

adopted is the use of a heterojunction between electron- and hole-accepting semiconductors, first demonstrated by Tang using a layered structure of phthalocyanine as hole acceptor and perylene derivative as electron acceptor [11]. The scheme of operation is shown in figure 5, with representative semiconducting polymers that also achieve this function [12].

Exciton ionisation using the heterojunction schemes shown in figure 9 can be engineered by selecting a sufficiently large band-edge offset so that the intra-chain exciton binding energy is overcome, and there are many examples of systems which achieve this now in the literature, which are reviewed in [13,14]. These systems are relatively inefficient when constructed as two-layer abrupt heterojunctions because photogenerated excitons must be able to reach this heterojunction in order to ionise, and their characteristic diffusion range is generally only 10% or so of the thickness of material required for full optical absorption [15].

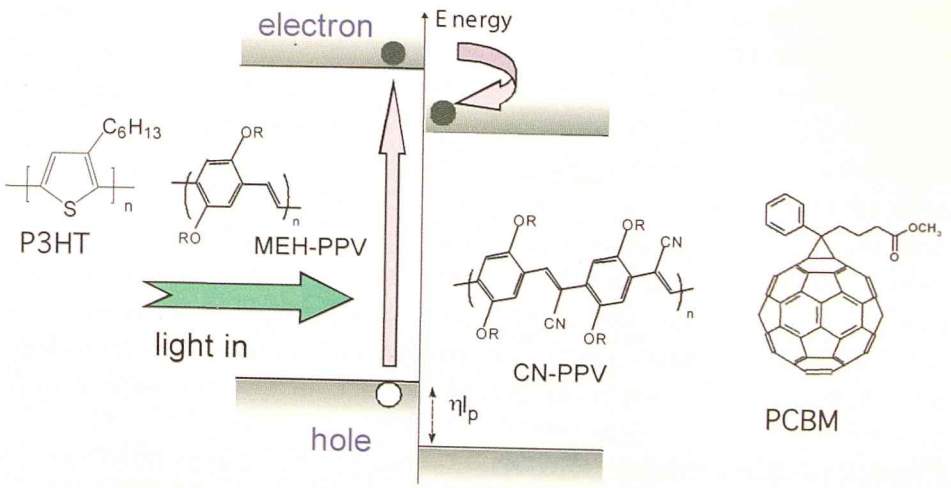


Figure 5: Energy band scheme for operation of a PV diode showing the steps of optical absorption in the hole-accepting polymer, followed by electron transfer at the heterojunction to the electron acceptor. This is demonstrated for alkoxy-substituted PPV versus a cyano-derivative of PPV [12] and for poly(3-hexyl thiophene) versus a soluble C60 derivative, PCBM [16].

One promising approach is to form a 'distributed heterojunction' structure as illustrated in figure 6, so that all photogenerated excitons are sufficiently close to heterojunctions to be ionised. With organic semiconductors we demonstrated this by the use of de-mixing of a blend of the electon- and hole-accepting materials upon evaporation of solvent from a spin-coated film of a mixture of the two [12]. The required length scale for the network is set by the diffusion range of photogenerated excitons, and this is of the order of 10 nm [15]. There are many refinements that would be desirable to use in order to control the morphology of the resultant network, for example, the use of diblock copolymers, but the chemical synthesis of such materials remains challenging.

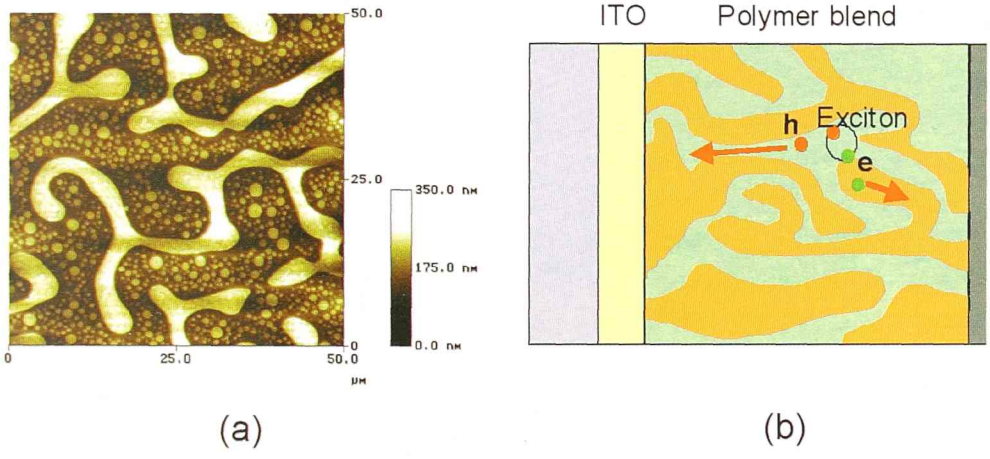


Figure 6. (a) Atomic force microscope image of the free surface of a blend of two semiconducting polymers that have partly de-mixed on evaporation of the solvent used to deposit a film of a mixture of the two by spin coating, after [17]; (b) idealised scheme for the distributed heterojunction required to achieve efficient charge photogeneration.

The best results to date have been achieved with blends of poly(3-hexyl thiophene) and the soluble derivative of fullerene, PCBM, as shown in figure 9. Energy conversion efficiencies at 5% are now reliably reported by several groups, as reviewed by [14]. This efficiency is limited by the relatively high onset of optical absorption, near 650 nm, for P3HT (PCBM is weakly absorbing in the visible and near-IR) and the relatively low open circuit voltage (typically 800 mV) due to the very substantial offset of conduction band energies between the two of almost 1 V. Both are in principle open to

engineering and improvement, but have yet to be easily improved.

Field-Effect Transistors

We demonstrated that it was possible to fabricate well-performing field-effect transistors, FETs, in the late 1980s using the semiconducting polymer polyacetylene [2], and these are now routinely fabricated using a wide range of organic semiconductors. Surface charge density is induced at the interface between organic semiconductor and a suitable dielectric by applying an electric field across the dielectric layer (through control of the gate voltage), and this induced charge density provides the conduction path between source and drain. Doped silicon wafers with thermally-grown oxide are often used to provide the gate and dielectric layer, with source and drain provided by lithographically-patterned gold (this provides a good choice for p-type operation). This structure only requires deposition of organic semiconductor (by sublimation or by solution-processing) to complete the transistor, and is therefore particularly convenient to work with.

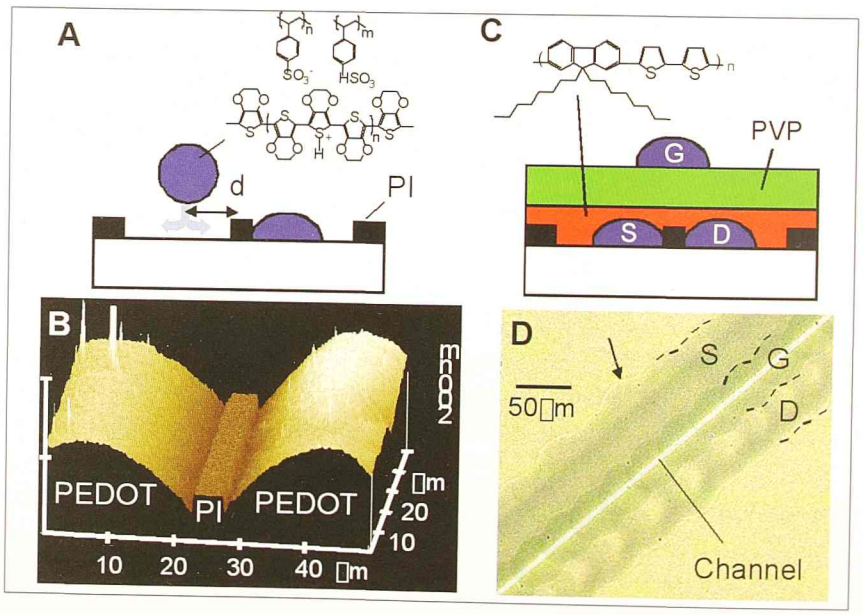


Figure 7 A shows ink-jet printing of PEDOT-PSS onto a surface-patterned substrate, causing confinement of the source and drain to either side of a 5 μm channel of polyimide, B shows an AFM image of the structure, C shows the completion of the device by spin-coating of semiconductor and dielectric and ink-jet printing of the gate electrode [18].

These structures are useful for exploration of the underlying semiconductor physics, but are not directly useful in areas requiring large areas. A greater challenge is to construct the transistor with organic materials for semiconductor, metallic electrodes and dielectric; if this can be demonstrated to be practical then it becomes feasible to manufacture arrays of transistors over large areas. Figure 7 shows the structure of such a scheme [19] that uses an organic polymer, poly(vinylphenol) as dielectric, and is patterned through the use of inkjet printing of PEDOT-PSS for the source, drain and gate. These devices show modest performance as FETs, as shown in the characteristics presented in figure 8, with field-effect mobilities of order 1012 cm2/Vs.

The mobilities of the field-effect charges control the usefulness of the device both for research and for applications. The challenge for materials process is to achieve a well-ordered interface between semiconductor and dielectric, because the surface-induced charges are confined within a couple of nm to this interface. Chemical treatment of the SiO2 top surface, for example with an alkyl silyl chloride to replace the polar SiO2 surface with a sheet of alkyl chains to the organic semiconductor layer, can greatly improve this order. By such methods, we reported mobilities for the best ordered polymers (P3HT) of up to 0.1 to 0.2 cm2/Vs [20]. Modified polythiophene structures that are better able to crystallise, PBTTT as illustrated in figure 13 below, have recently been shown to exhibit considerably improved mobilities, up to 0.5 cm2/Vs [21].

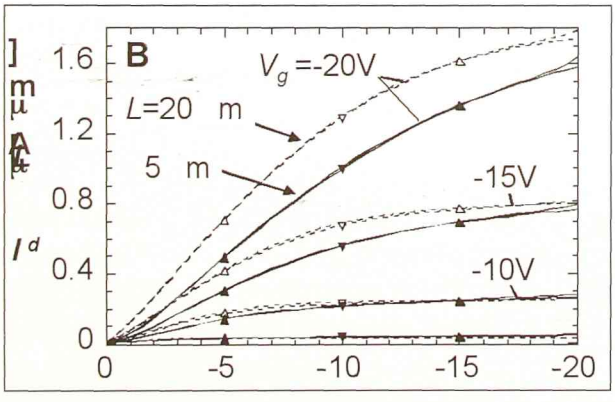


Figure 8: Drain current x channel length versus source-drain voltage for transistors as shown in figure 7, for the values of gate voltage indicated on the figure. Note that there is evidence for some contact resistance, more noticeable for the shorter channel structure [19]

Light-Emitting Transistors

Most of the literature on organic semiconductor FETs has reported p-type operation of these devices, partly because it is generally more convenient to work with air-stable electrodes such as gold, but primarily because n-type operation is difficult to observe, particularly for higher energy gap semiconductors. However, many of the problems in observing n-type behaviour can be attributed to the interfacial chemistry that can occur when polar dielectrics are used, particularly so for the case of SiO2, for which electrochemical reactions to reduce surface hydroxyl groups to O- ions can prevent injection of electrons into the semiconductor conduction band [22]. When non-polar organic dielectrics are used, n-type operation and ambipolar operation become relatively easy to observe [22]. This has made it possible to construct FETs which allow simultaneous injection of electrons (from 'source') and holes (from 'drain') into the transistor channel, and which show light emission where the drifting electrons and holes meet within the channel [23]. This is seen directly in images of light emission from within the transistor channel, as shown in figure 9.

FET Applications

Organic transistor at present show moderate performance in comparison with amorphous silicon as used (widely) in large-area thin-film transistor arrays for active matrix addressing of liquid crystal displays. Their potential advantage lies in the scope for low temperature processing, enabling the manufacture onto plastic substrates that cannot be taken to the high process temperatures required for amorphous silicon. We are using these for the manufacture of active matrix transistor backplanes onto plastic substrates for electronic paper displays [24].

Outlook

The past two decades were characterized by tremendous progress in the performance of organic semiconductor devices. The progress has been so remarkable, that the field of organic electronics – a electronics technology based on organic materials as active layers in semiconductor devices – has moved from the realm of academic laboratories into large-scale manufacturing. With organic LEDs appearing into consumer electronic devices, organic FETs close to the market, and the potential for progress with PV diodes, the strong technological basis being established makes the future of organic electronics seem bright.

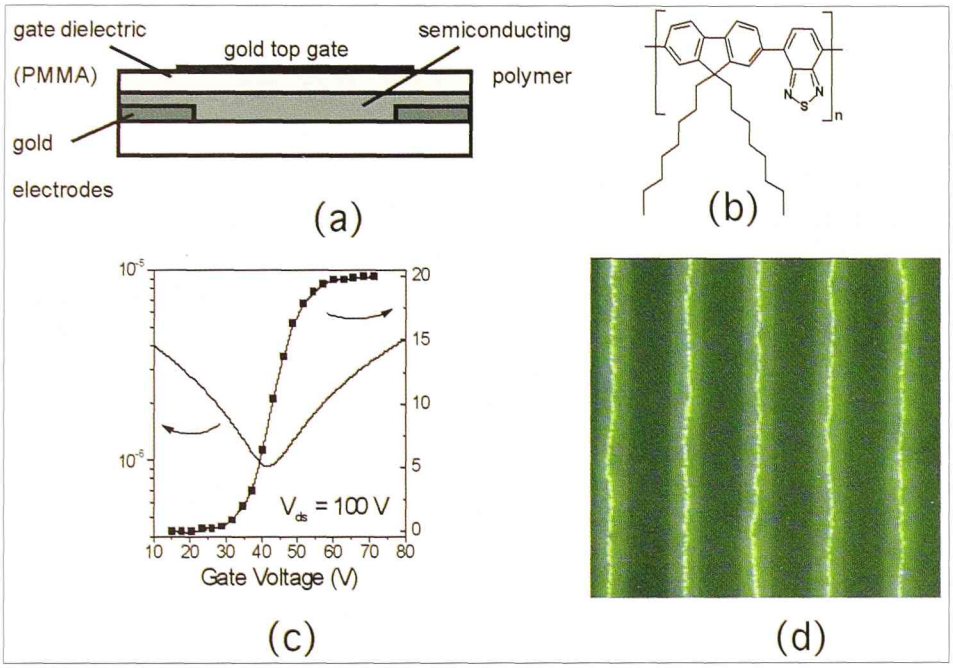


Figure 9 (a) Structure of a light-emitting transistor; (b) structure of the light-emitting polymer, F8BT; (c) position of emission zone with respect to source (grounded) electrode and source-drain current versus gate voltage during a transfer scan of an ambipolar F8BT transistor with L = 20 μm; (d) optical images of light emission from F8BT light-emitting transistor with interdigitated source-drain electrodes (dark areas) with L = 20 μm. [23].

References:

[1] C. W. Tang, S. A. VanSlyke: Appl. Phys. Lett. Vol. 51 (1987) p. 913-15.

[2] J. H. Burroughes, C. A. Jones, R. H. Friend: Nature Vol. 335 (1988) p. 137-41.

[3] J. H. Burroughes, D. D. C. Bradley, A. R. Brown, R. N. Marks, K. Mackay, R. H. Friend, P. L. Burns, A. B. Holmes: Nature Vol. 347 (1990) p. 539-41.

[4] M. Pope, C. E. Swenberg: Electronic Processes in Organic Crystals and Polymers, Oxford University Press, New York, 1999.

[5] M. Pope, H. Kallmann, P. Magnante: J. Chem. Phys. Vol. 38 (1963) p. 2042-43.

[6] W. Helfrich, W. G. Schneider: Phys. Rev. Lett. Vol. 14 (1965) p. 229-31.

[7] C. K. Chiang, C. R. Fincher, Y. W. Park, A. J. Heeger, H. Shirakawa, E. J. Louis, S. C. Gau, A. G. MacDiarmid: Phys. Rev. Lett. Vol. 39 (1977) p. 1098-101.

[8] H. Sirringhaus, T. Shimoda: Materials Research Society Bulletin, Vol. 28, No. 11, November Vol. (2003).

[9] C. W. Sele, T. von Werne, R. H. Friend, H. Sirringhaus: Advanced Materials Vol. 17 (2005) p. 997-+.

[10] www.novaled.com.

[11] C. W. Tang: Appl. Phys. Lett. Vol. 48 (1986) p. 183-85.

[12] J. J. M. Halls, C. A. Walsh, N. C. Greenham, E. A. Marseglia, R. H. Friend, S. C. Moratti, A. B. Holmes: Nature Vol. 376 (1995) p. 498-500.

[13] B. A. Gregg: Journal of Physical Chemistry B Vol. 107 (2003) p. 4688-98.

[14] S. Gunes, H. Neugebauer, N. S. Sariciftci: Chemical Reviews Vol. 107 (2007) p. 1324-38.

[15] J. J. M. Halls, K. Pichler, R. H. Friend, S. C. Moratti, A. B. Holmes: Appl. Phys. Lett. Vol. 68 (1996) p. 3120-22.

[16] G. Yu, J. Gao, J. C. Hummelen, F. Wudl, A. J. Heeger: Science Vol. 270 (1995) p. 1789-91.

[17] A. C. Arias, J. D. MacKenzie, R. Stevenson, J. J. M. Halls, M. Inbasekaran, E. P. Woo, D. Richards, R. H. Friend: Macromolecules Vol. 34 (2001) p. 6005-13.

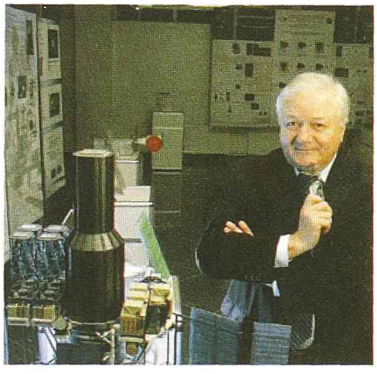
[18] J. Zaumseil, R. H. Friend, H. Sirringhaus: Nature Materials Vol. 5 (2006) p. 69-74.
[19] H. Sirringhaus, T. Kawase, R. H. Friend, T. Shimoda, M. Inbasekaran, W. Wu, E. P. Woo: Science Vol. 290 (2000) p. 2123-36.
[20] H. Sirringhaus, P. J. Brown, R. H. Friend, M. M. Nielsen, K. Bechgaard, B. M. W. Langeveld-Voss, A. J. H. Spiering, R. A. J. Janssen, E. W. Meijer, P. Herwig, D. M. de Leeuw: Nature Vol. 401 (1999) p. 685-88.
[21] I. McCulloch, M. Heeney, C. Bailey, K. Genevicius, I. Macdonald, M. Shkunov, D. Sparrowe, S. Tierney, R. Wagner, W. M. Zhang, M. L. Chabinyc, R. J. Kline, M. D. McGehee, M. F. Toney: Nature Materials Vol. 5 (2006) p. 328-33.
[22] L. L. Chua, J. Zaumseil, J. F. Chang, E. C. W. Ou, P. K. H. Ho, H. Sirringhaus, R. H. Friend: Nature Vol. 434 (2005) p. 194-99.
[23] J. Zaumseil, C. R. McNeill, M. Bird, D. L. Smith, P. P. Ruden, M. Roberts, M. J. McKiernan, R. H. Friend, H. Sirringhaus: Journal of Applied Physics Vol. 103 (2008).
[24] www.plasticlogic.com: Vol.

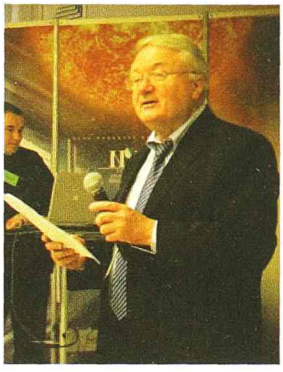
Professor Rashid A. Sunyaev

Co-Winner of the 2009 King Faisal International Prize for Science

Synopsis of Achievements

Professor Rashid Sunyaev is a prominent Russian physicist whose outstanding contributions to high energy astrophysics and cosmology have profoundly impacted both fields and placed him at the forefront of contemporary astrophysicists. Born in Tashkent, Usbekistan on March 1, 1943, Professor Sunyaev graduated from Moscow Institute of Physics and Technology in 1966 and received his Candidate of Sciences (PhD equivalent) and Doctor of Sciences degrees from Moscow University in 1968 and 1973, respectively. Between 1968-1974, he served as a scientific researcher at the Institute of Applied Mathematics and subsequently as Head of the Laboratory of Theoretical Astrophysics at the Space Research Institute of the USSR Academy of Sciences in Moscow. He became full professor at Moscow Institute of Physics and Technology from 1975-2001 and Head of the High Energy Astrophysics Department of the Space Research Institute in Moscow from 1982-2002. He is currently Director of the Max Planck Institute for Astrophysics and Chief Scientist at the

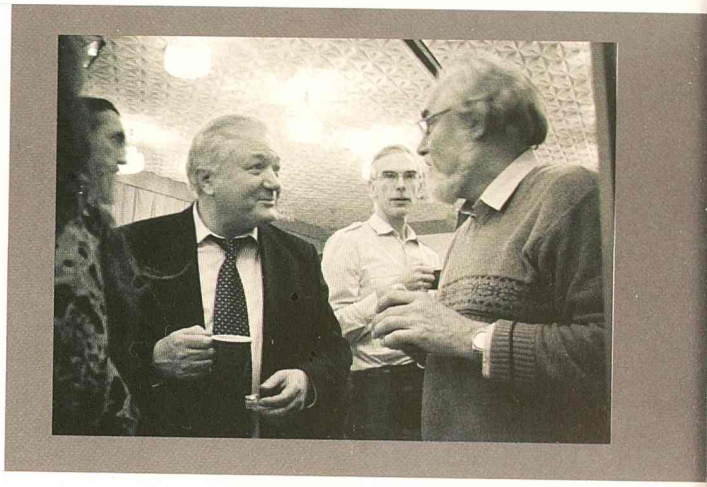




Russian Space Research Institute He is also Russia's principal scientific investigator of the International Gamma Ray Astrophysics Laboratory (INTEGRAL) of the European Space Agency.

Professor Sunyaev's fundamental contributions to the advancement of cosmology and astrophysics during the past thirty years cannot be over-emphasized. Among his most distinguished contributions are: his predictions of acoustic peaks in the cosmic microwave background angular distribution, and the development of both the Sunyaev-Zeldovich effect (S-Z effect) on clusters of galaxy and the theory of disk accretion (Standard Shakura-Sunyaev disk) and observational appearance of black holes in binary systems and active galactic nuclei. These and several other achievements by Sunyaev have driven theoretical developments to new frontiers and led to the generation of powerful and widely used tools to study structures in the universe. Sunyaev has also made significant contributions to space science. He has led the team that built the X-ray observatory on Mir space station and the GRANAT orbiting X-ray observatory and is currently working with his team in preparing the world's first astronomical X-ray satellite and on other projects related to the Plank Mission of the European Space Agency. .

Professor Sunyaev's outstanding accomplishments have been recognized by numerous honors and awards. He is a fellow of the US National Academy of Science, the Russian Academy of Science and the



Royal Netherlands Academy of Arts and Sciences, and an honorary member of Bashkortostan and Tatarstan Academies of Sciences. He is also a member of the International Astronomical Union and former vice-president of its space research committee, member and former vice-president of the European Astronomical Society, member of the American Physical Society, international member of the American Philosophical Society and foreign fellow of the Royal Astronomical Society. In addition, Professor Sunyaev has held numerous visiting and honorary professorships, Lectureships and visiting scientist/ scholar positions at leading universities including Johns Hopkins University, Columbia University, University of California at Berkley,

University of Virginia, Harvard-Smithonian Center for Astrophysics, Institute for Advanced Studies at Princeton University, California Institute of Technology, Cambridge University, Massachusetts Institute of Technology, Ludwig-Maxmillian University, Leiden University, Toronto University and the Bose National Center for Basic Sciences in Calcutta. He is also the recipient of several prestigious awards including the Crafoord Prize of the Royal Swedish Academy of Sciences, the Heinemann Prize for Astrophysics of the American Physical Society and the Gold Medal of the Royal Astronomical Society of the UK.

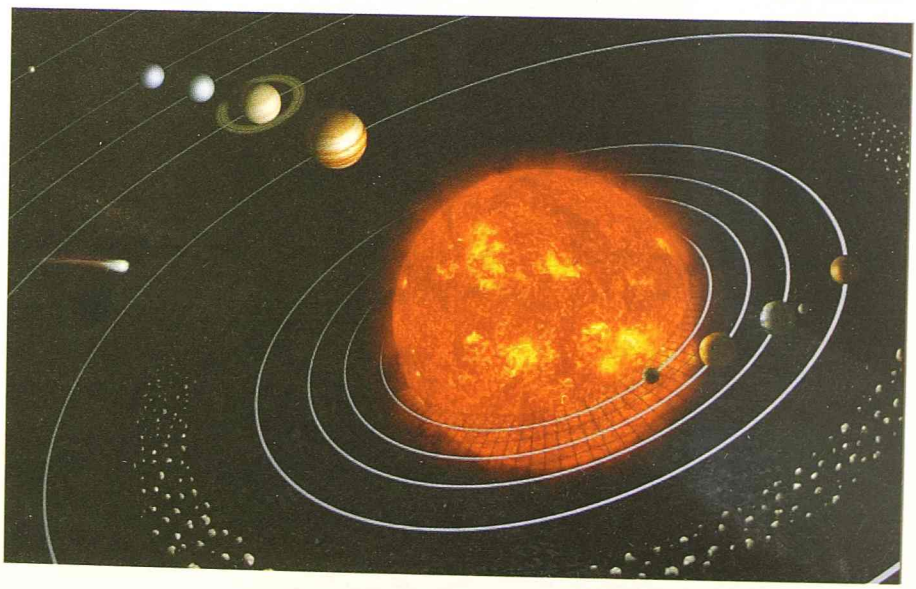
Professor Sunyaev has published over 300 papers, some of which stand out among the most highly cited publications in astrophysics.

Clusters of galaxies and Cosmic Microwave Background Radiation

Rashid A. Sunyaev
Max-Planck-Institute for Astrophysics
Karl-Schwarzschid-Strasse 1
85741 Garching bei Muenchen - Germany

We are living in the Solar System, which consists of a very ordinary star – the Sun - with planets rotating around it. Our Earth is certainly the best among them, because this is our home.

Our Sun is one of many billions of stars we observe in the Milky Way – our Galaxy.



Fortunately for us the Sun is located rather far from the Center of our Galaxy, where the super massive black hole is located. Its mass exceeds the mass of our Sun four million times.

The total mass of the Galaxy exceeds the mass of our Sun a few hundred billion times.

Billions of galaxies are observable with modern ground based and space telescopes. Some of them form **clusters of galaxies**, containing hundreds and even thousands of galaxies.

The mass of these objects is correspondently high, exceeding the mass of our Galaxy thousands of times. Galaxies are moving in the common

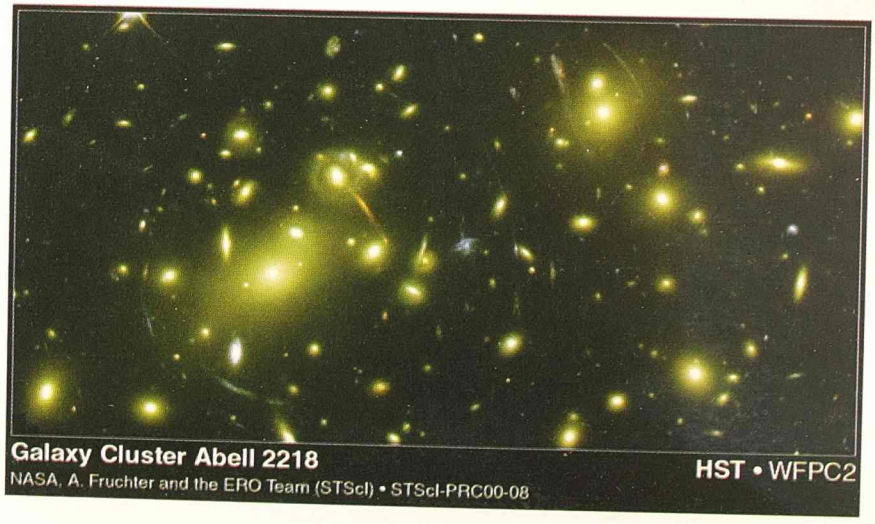
gravitational potential of these clusters with velocities of the order of 1000 km/second. This gravitational potential is so high that the images of some more distant galaxies beyond the cluster are distorted and their surface brightness is strongly amplified due to gravitational lensing. The physics of clusters of galaxies and the ways to observe the hot gas inside them will be the main subject of my presentation.

In 1970 the X-Ray emission from nearby clusters of galaxies was detected. There were two possible interpretations of this discovery –

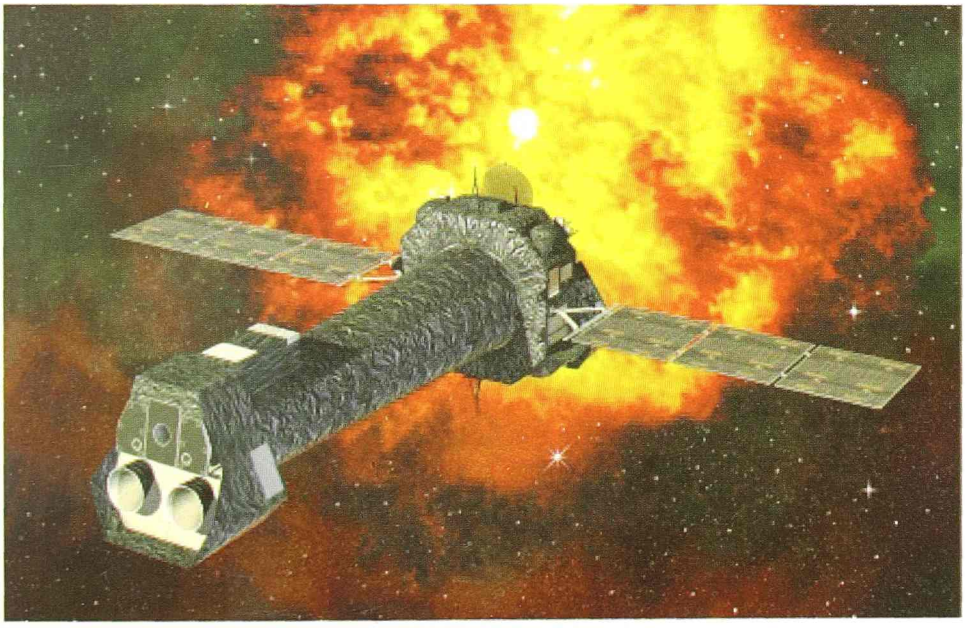
a) X-Rays could be produced due to collisions of fast electrons with protons in the hot intergalactic gas between galaxies. To create an extended atmosphere around cluster of galaxies gas should be very hot with a temperature Te of the order of tens of millions or even hundreds of millions of degrees. The speed of sound in such a gas is close to 1000 km/s. This means that galaxies are moving within the gas with velocities close to the speed of sound.

b) Observed X-Rays could be produced due to the interaction of the ultrarelativistic electrons and radiophotons, emitted by the same electrons moving in the magnetic field due to the synchrotron radiation mechanism.

It was necessary to find which physical mechanism is working. Especially interesting was to find a way to perform the necessary observations from the ground because X-Rays are completely absorbed by the Earth's atmosphere and it is necessary to launch spacecraft to observe the sky in X-Rays.



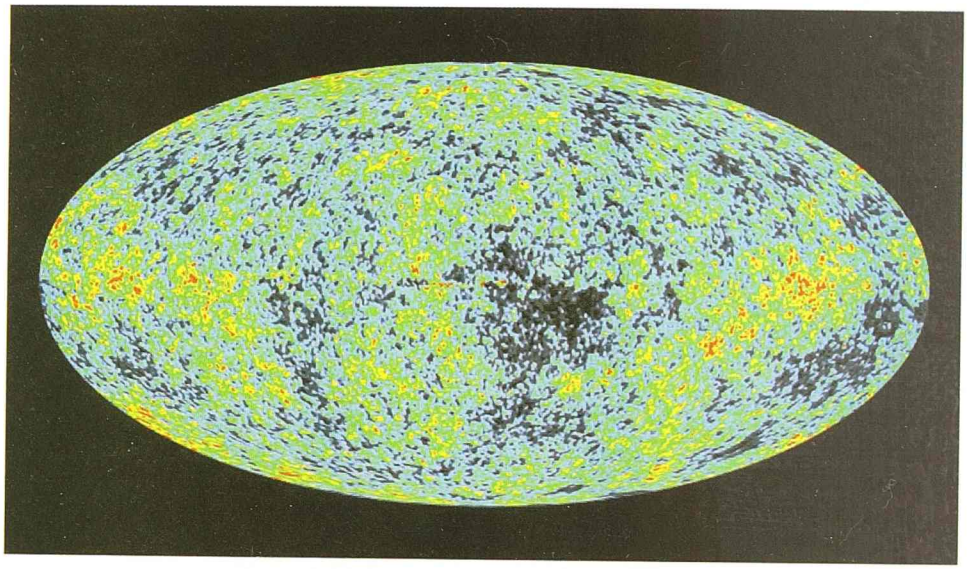
I will describe below the method we proposed for this purpose in 1970. But before going into the details of our method I should mention that in 1965 the future Nobelists, Arno Penzias and Robert Wilson, discovered Cosmic Microwave Background Radiation (CMB), which was predicted by George Gamov in the beginning of the 1950s. This radiation has a black body spectrum corresponding to a temperature of 2.7 degrees on the Kelvin scale, and every cubic cm of space in the Universe contains today 400 CMB photons. The majority of these photons are in the radio spectral band or in the submillimeter band. This radiation is extremely isotropic.



The WMAP spacecraft obtained the microwave sky map with unprecedented precision. The observations took 5 years. The brightest (red) spots on the sky map correspond to 250 microKelvin increase of the radiation brightness temperature, the dimmest (blue) spots have a brightness temperature of 250 microKelvin lower than the average 2.7 K. The angular fluctuations of the CMB radiation field are tiny – at least ten thousand times smaller than the total radiation flux.

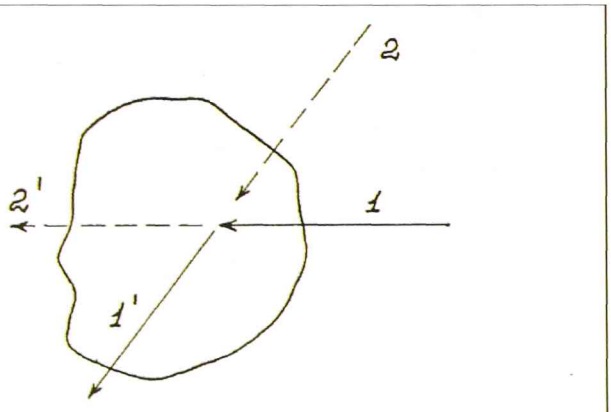
Every cluster of galaxies with hot intergalactic gas inside it is illuminated from all sides by CMB photons. In 1971 I started to think – does the interaction of CMB photons with the hot electron of the

intergalactic gas produce any detectable traces on CMB brightness in the direction toward clusters of galaxies?



First attempt – just to check the screening of CMB sky brightness due to the finite optical depth of the cluster of galaxies gave a negative result. Photon 1 was scattering out in the direction towards the observer, but in the isotropic radiation field another photon 2 will replace it after scattering onto another electron. At that time (in 1970) my mentor Prof. Zeldovich (his name is well known due to his great results in physics, the theory of combustion, ecology, the theory of shock waves, cosmology and for his giant contribution to the creation of the Soviet nuclear and hydrogen bombs) and I found analytical solutions describing the exchange of energy between plasma and radiation in the early Universe.

I tried to use the same approach to the situation in clusters of galaxies. The simplest way to explain the process follows. Let us consider the spectrally narrow laser line illuminating the cloud of thermal electrons from all sides and let us compute (or measure) the spectra of scattered photons. Let us consider the case of low optical depth when the probability of photons scattering inside cloud is rather low. Then we can easily count all the photons that experienced only one scattering on electrons inside the cloud.

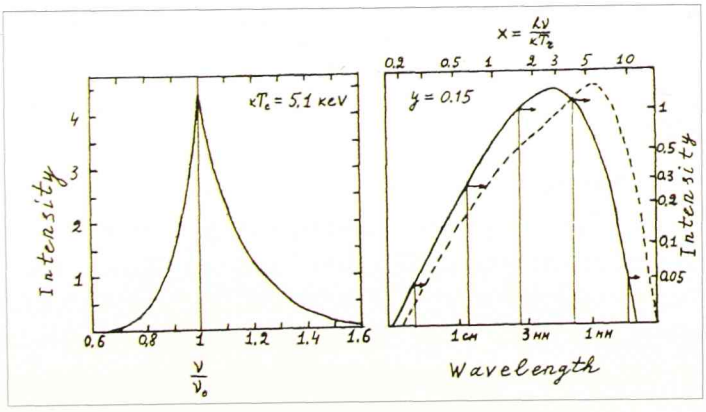


The scattering of isotropic radiation field by the cloud of electrons.

The resulting spectrum is presented in the left image in the figure below. The laser line is broadened due to the stochastic nature of the motion of thermal electrons. Their average velocity at the electron temperature of 5.1 KeV or 55 million degrees is very high (of the order of 1/7 of the speed of light). To get correct results we should take into account the effects of the special theory of relativity. The result is very interesting: the right (high energy) wing of the line is stronger than the left one. This is equivalent to the effective shift of the line toward higher frequencies (hot electrons give part of their energy to low frequency photons. The plasma cools down and the radiation field increases its energy density).

SCATTERING OF RADIATION BY HOT MAXWELLIAN ELECTRONS

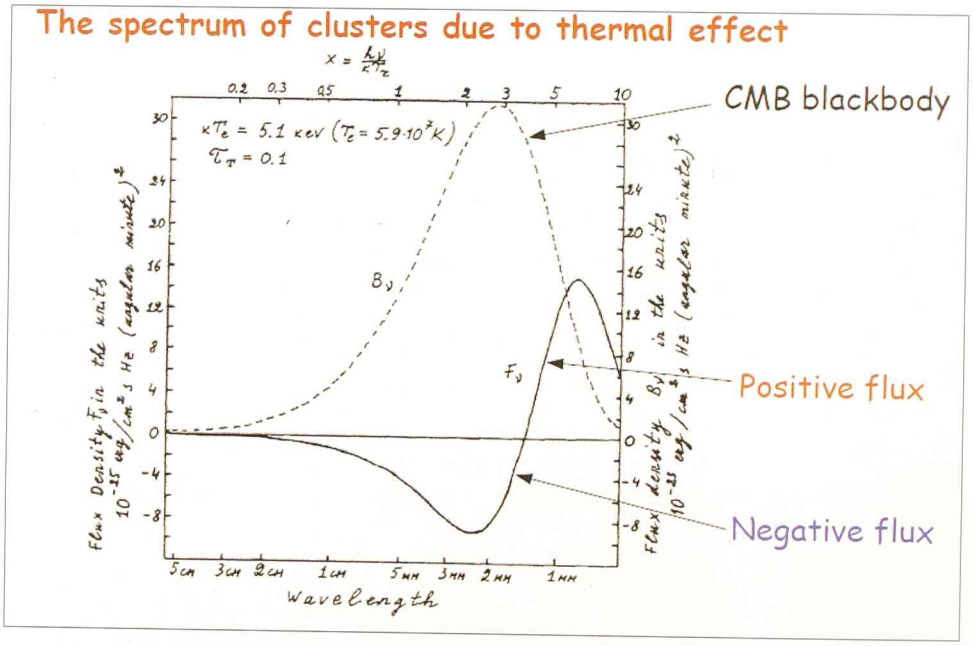
spectral changes due to doppler-effect on moving electrons with kTe~ 5 KeV and average velocity of the order of 1/7 c



Laser line

Line is broadened and effectively shifted toward higher frequencies due to second order effects in v/c

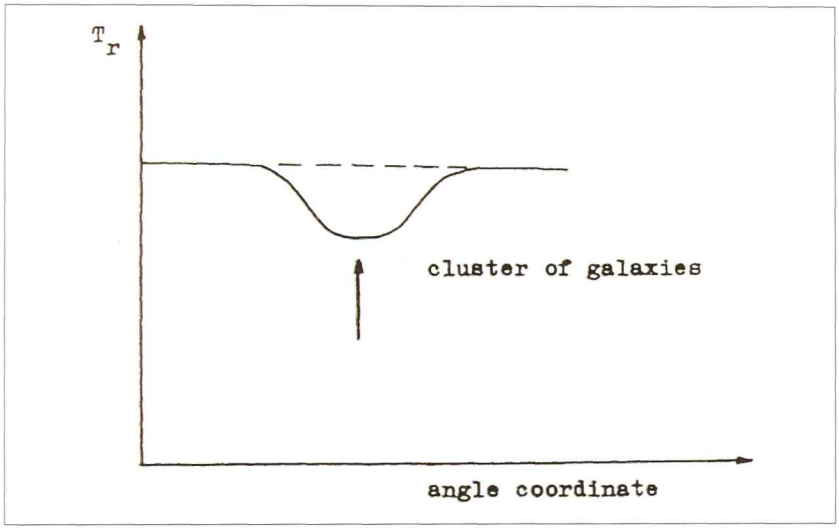
If we will represent the continuum black body spectrum as a sum of many narrow lines, we will see how the resulting radiation spectrum will form (the right hand plot on the figure above). Photons will move towards higher frequencies and the spectrum of CMB in the direction of the cluster with hot gas will change: the radiation brightness at low



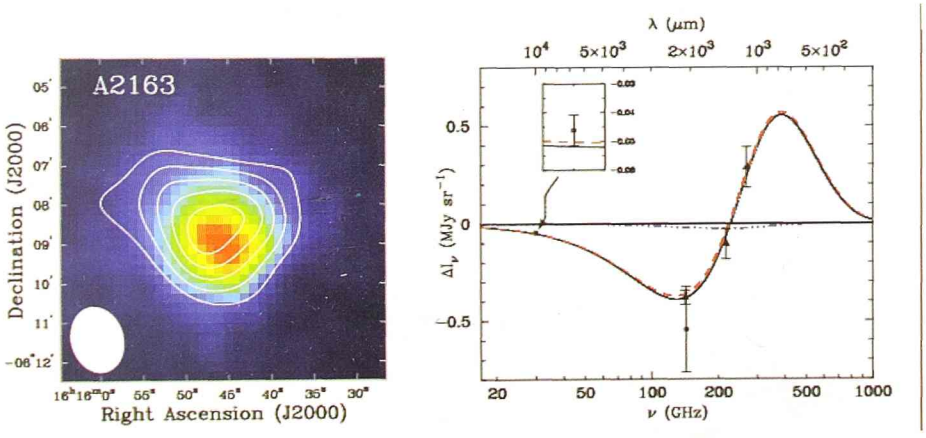
frequencies (cm and mm band) will decrease but it will increase at higher frequencies.

As a result in the centimeter spectral band clusters should be observed as a ***holes*** in the sky average brightness defined by CMB intensity.

It was proposed to look for such "negative" sources in the sky. The same regions with the hot gas should be bright "positive" sources at higher frequencies. The prediction of "negative" sources in the directions to the clusters with hot gas was very unusual. According to our normal experience "hotter" objects are brighter than "cooler" objects. However the prediction was correct because beautiful and simple physics was behind it.



This prediction was made in 1970. Several experimental groups were working for fifteen-twenty years to detect the effect. Finally in the second half of 1980's they were able to detect the predicted effect in the direction of several well known clusters of galaxies. Very soon spectral observations confirmed that.

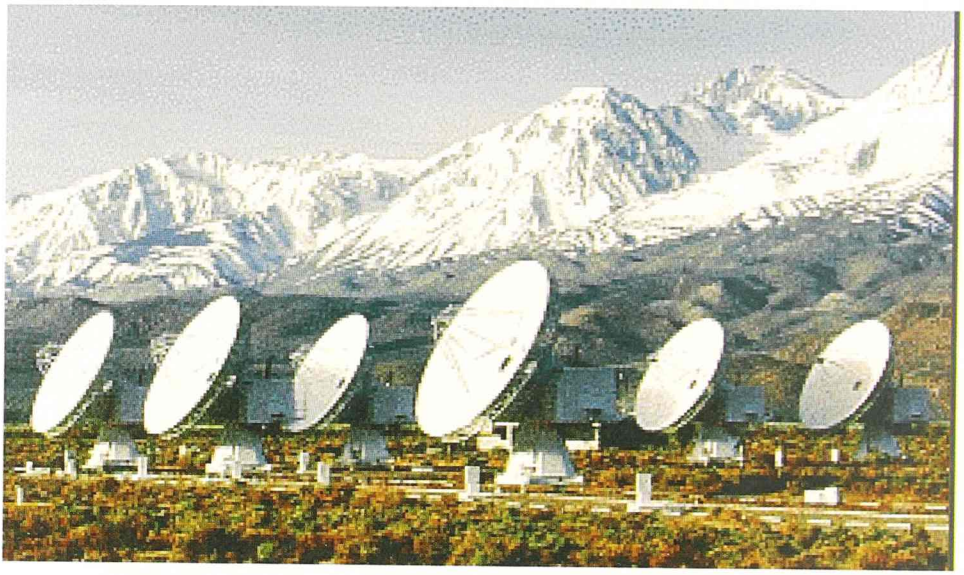


"negative" sources transform to "positive" at higher frequencies.

Here I am glad to write words of gratitude to the radio astronomers who spent a significant part of their lives to detect the thermal SZ-effect and to make it possible to use these measurements for the purposes of

cosmology. I should mention John Carlstrom, Mark Birkinshaw and Francesco Melchiorri first.

The interferometric observations of John Carlstrom's group (University of Chicago) confirmed one of the most important predictions of the theory:

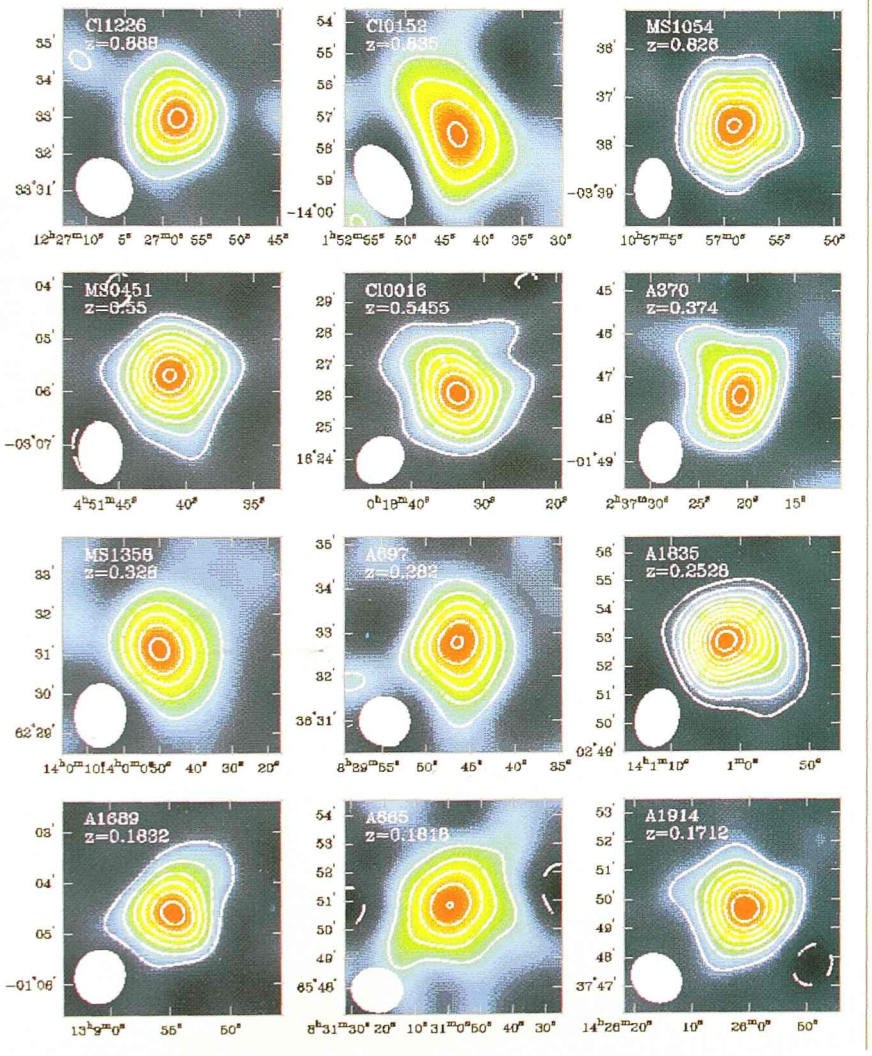


the radiation spectrum and microwave brightness of the cluster does not depend on the distance to the cluster (or its redshift z). Only the angular dimensions of the cluster in a microwave spectral band depend on the distance. This means that we have a possibility to detect all the clusters of galaxies in the observable Universe.

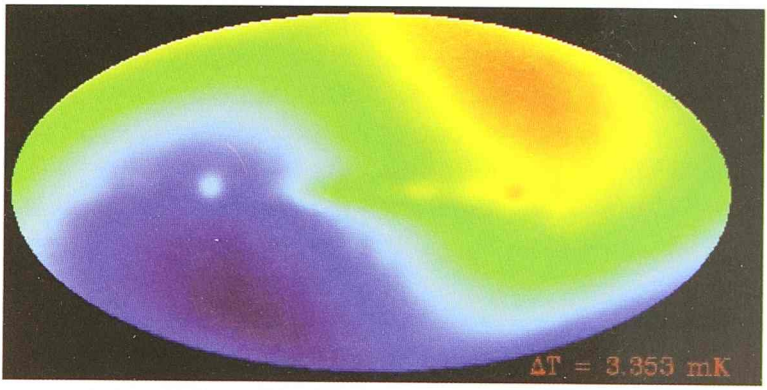
a) Observations of clusters in a microwave spectral band and X-Rays permit us to measure the distances to the clusters of galaxies. The distance scale is very important for cosmology and our understanding of the properties of our Universe.

b) The counting of distant clusters of galaxies will permit us to find how numerous they are at different distances. This function depends strongly on the main parameters of our Universe – its density, composition etc. We hope that observations of clusters will permit us to get additional information about the properties of dark matter and dark energy.

c) Observations at three frequencies (< 160 GHz), near cross-point at 217 GHz and at high frequencies (~ 270 GHz) will permit us to measure the peculiar velocities of clusters of galaxies at any distance. We know that CMB is isotropic only in one coordinate frame. If we are moving relative that frame an additional (dipole) component appears in the CMB angular distribution. Such a strong dipole component exists in the CMB maps observed for example from the COBE spacecraft . The sky is brighter in the direction of the Sun motion and dimmer in the opposite direction.



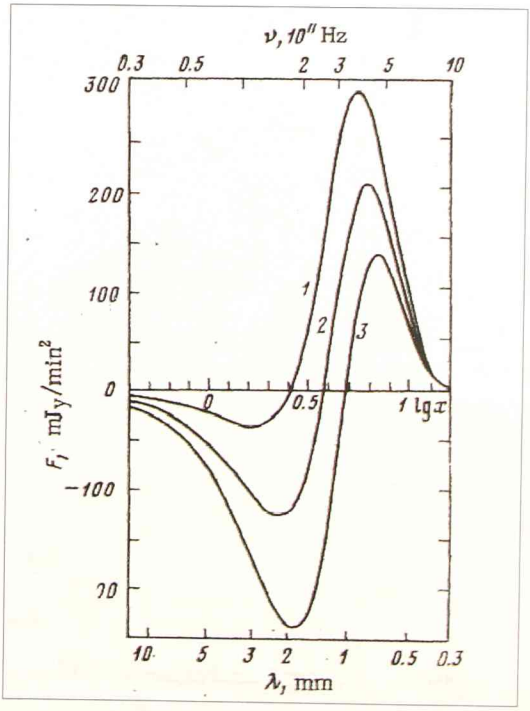
Now astronomers are trying to use the SZ-effect for cosmology.



This change of brightness corresponds to the velocity equal to 630 km/sec. Only part of this value is due to the rotation of our Galaxy.

The WMAP spacecraft detects annual changes in the sky's brightness connected with Earth's motion around the Sun (30 km/sec) due to the same reason.

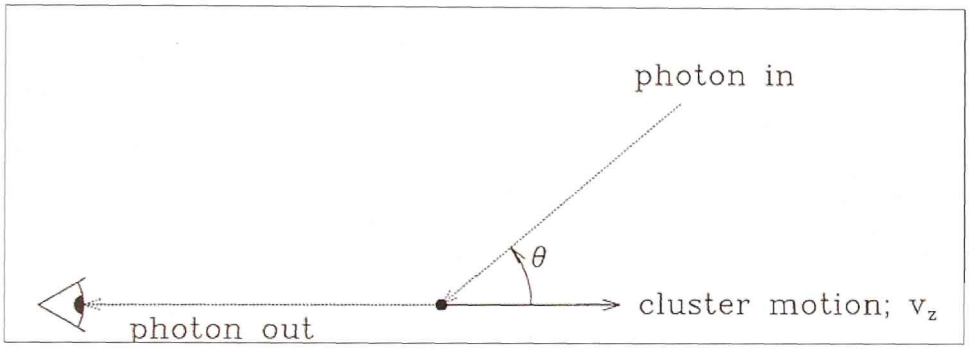
Hot gas in the distant clusters could serve as an observer measuring the isotropy of CMB in the place of its location and therefore permits us to measure the velocity of the cluster relative to CMB.



Kinetic SZ-effect

The figure above demonstrates how the spectrum of the SZ-effect depends on the peculiar velocity of the cluster: 1) Vz = - 3000 km/sec, 2) Vz = 0; 3) Vz = 3000 km/sec (a plus sign indicates the receding cluster).

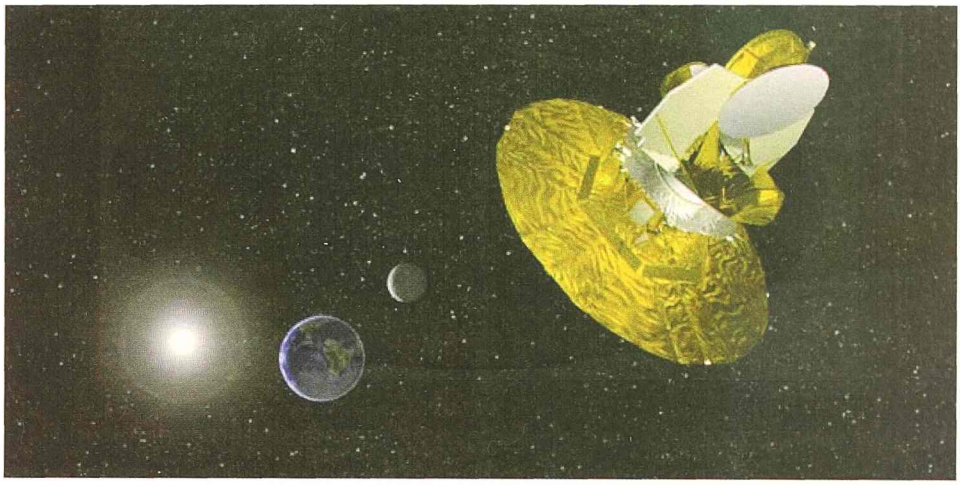
In the second half of this April the European Space Agency is planning to launch the PLANCK Surveyor spacecraft into the second Lagrangian



$$\frac{\Delta T_{\text{rad}}}{T_{\text{rad}}} \approx -\tau_e \frac{v_z}{c}$$

NASA's WMAP spacecraft has already been working for more than 7 years near the Second Lagrangian Point. This position is great – the Sun, Moon and the Earth are all time on the same side of the spacecraft. This opens a possibility to use very effectively the cryogenic bolometer detectors of PLANCK mission which will work at a temperature of only 100 milliKelvin above absolute zero. One of the goals of the PLANCK

point of the Solar System, which is 1,5 million km away from the Earth.



mission is to map thousands of clusters of galaxies on the whole sky using the SZ-effect. It is planned to measure the peculiar velocities for many bright clusters.

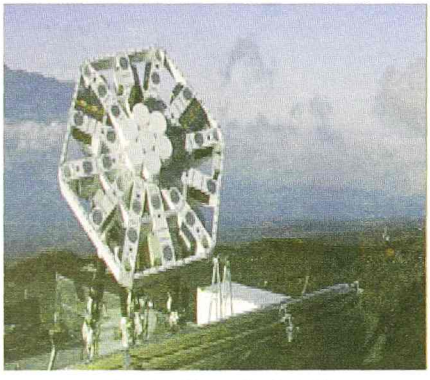
The PLANCK spacecraft will compete with several ground based telescopes specially constructed to look for unknown clusters of galaxies on the sky. These instruments have better angular resolution than PLANCK, have more sophisticated detectors but the thermal radiation of the Earth's atmosphere makes these observations difficult. Therefore radio astronomers choose for these telescopes the dryest and highest

2800 m altitude, 980 bolo meters

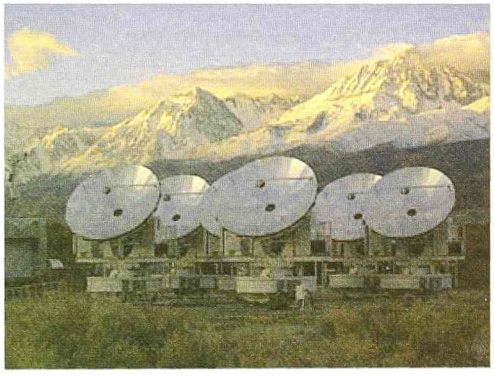
Site of South Pole Telescope
Now 10m telescope is staying there

Atacama Cosmology Telescope
5 200 m height, 3024 bolometers
SZ Array, California

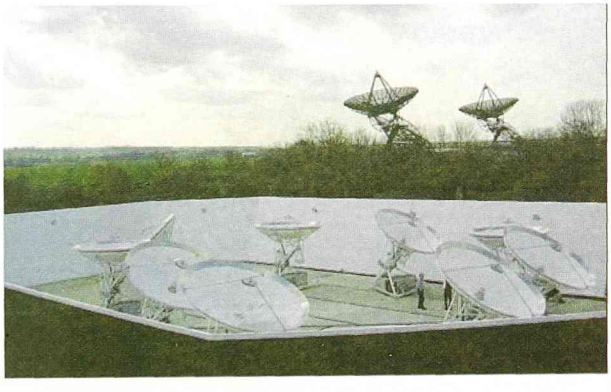
APEX, 12 m,330 bolometers, Chile,5000m



AMIBA, Hawaii, 3,396 m



SZ Array, California

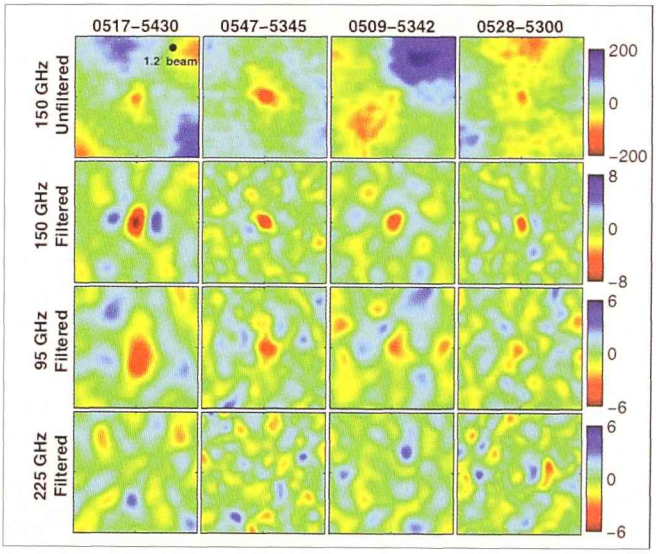


AMI, Cambridge

places on the globe like the South Pole in Antarctica or Atacama Desert in Chile with a site at an altitude of 5000 m. However observations of AMI at lower frequencies (13-18 GHz) are going successfully even in foggy British Cambridge, because atmosphere is much more transparent at those

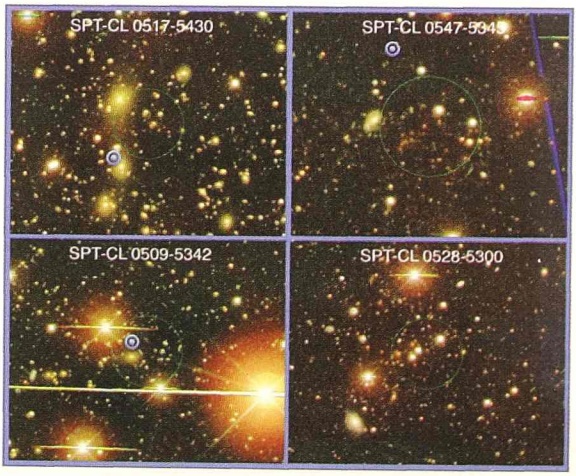
frequencies.

I can show you here the recent discovery of four very distant and rich, previously unknown, clusters of galaxies by the South Pole Telescope (John Carlstrom's group). They all were detected as "negative" sources



at 95 and 150 GHz and they are "absent" at the crossover frequency 225 GHz (as predicted by theory).

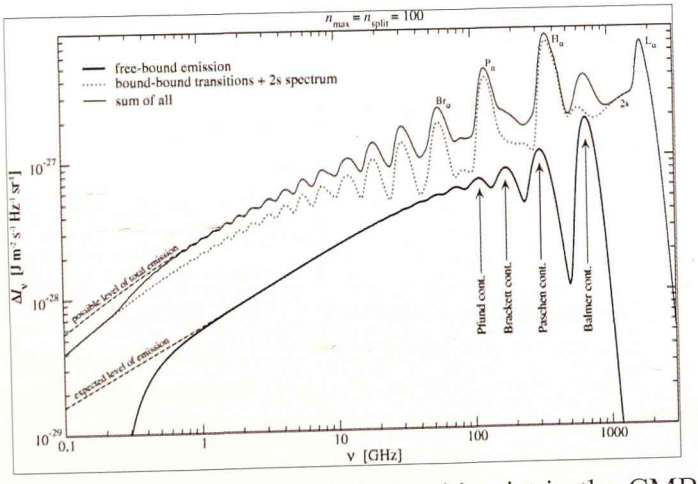
Optical observations immediately confirmed the existence of rich clusters of galaxies in the same positions and gave estimates of the distance to these clusters.



The South Pole Telescope and the Atacama Cosmology Telescope continue their sky surveys. Both teams hope to detect thousands of unknown clusters of galaxies in the next several years.

It is an interesting feeling when you are young and have an extremely strong but very busy mentor (as I had), and are trying to propose the existence of the new features on our sky. The idea and predictions are very unusual and it is not easy to convince skeptics. In addition it was necessary to wait for a long time before this prediction and the prediction of "acoustic peaks" in the angular distribution of CMB (made together with Yakov Zeldovich also in 1970, and independently by Peebles and Yu the same year) were confirmed by observers using the most sophisticated spacecraft, detectors and telescopes. I do not have room in this short paper to show the figures describing the theory of acoustic peaks and to demonstrate to you the great results of the Boomerang and Maxima-II balloon flights and WMAP spacecraft, which proved their existence. It is enough for me to know that they have been found with great significance on our sky and they will remain there for billions of years.

I prefer to show you in the figure below our recent prediction made together with my younger colleagues Jens Chluba and Jose Alberto Rubino-Martin.



These tiny spectral features which should exist in the CMB spectrum. They represent the high level transitions in the hydrogen atoms which recombined long ago and the corresponding lines have redshifted 1500 times. I hope that these features will be discovered during the next 5 – 15 years and will tell us a lot about the history of our Universe.

Literature:

Sunyaev, R. A. ; Zeldovich, Ya. B. , Small-Scale Fluctuations of Relic Radiation, Astrophysics and Space Science, Vol. 7, pp.3-19, 1970

Peebles, P. J. E. ; Yu, J. T. , Primeval Adiabatic Perturbation in an Expanding Universe, Astrophysical Journal, vol. 162, p.815, 1970

Sunyaev, R. A. ; Zeldovich, Ya. B. , The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies,
Comments on Astrophysics and Space Physics, Vol. 4, p.173, 1972

Sunyaev, R. A. , The Microwave Background Radiation in the Direction Toward Clusters of Galaxies, Soviet Astronomy Letters, Vol. 6, p. 213, 1980

Sunyaev, R. A. ; Zeldovich, Ya. B. , Microwave background radiation as a probe of the contemporary structure and history of the universe, Annual Review of Astronomy and Astrophysics., Vol. 18, p. 537-560, 1980

Sunyaev, R. A. ; Zeldovich, Ya. B. , The velocity of clusters of galaxies relative to the microwave background - The possibility of its measurement, Royal Astronomical Society, Monthly Notices, vol. 190, p. 413-420, 1980

Kogut, A. ; Lineweaver, C. ; Smoot, G. F. ; Bennett, C. L. ; Banday, A. ; Boggess, N. W. ; Cheng, E. S. ; de Amici, G. ; Fixsen, D. J. ; Hinshaw, G.; Jackson, P. D. ; Janssen, M. ; Keegstra, P. ; Loewenstein, K. ; Lubin, P. ; Mather, J. C. ; Tenorio, L. ; Weiss, R. ; Wilkinson, D. T. ; Wright, E. L. , Dipole Anisotropy in the COBE Differential Microwave Radiometers First-Year Sky Maps, Astrophysical Journal, Vol. 419, p.1,1993

Birkinshaw, M. Physics Reports, The Sunyaev-Zel'dovich effect Vol. 310, p. 97-195, 1999

Carlstrom, John E. ; Holder, Gilbert P. ; Reese, Erik D. , Cosmology with the Sunyaev-Zeldovich Effect, Annual Review of Astronomy and Astrophysics, Vol. 40, p. 643-680, 2002.

Bennett, C. L. ; Halpern, M. ; Hinshaw, G. ; Jarosik, N. ; Kogut, A. ; Limon, M. ; Meyer, S. S. ; Page, L. ; Spergel, D. N. ; Tucker, G. S. ; Wollack, E. ; Wright, E. L. ; Barnes, C. ; Greason, M. R. ; Hill, R. S. ; Komatsu, E. ;Nolta, M. R. ; Odegard, N. ; Peiris, H. V. ; Verde, L. ; Weiland, J. L. , First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results, The Astrophysical Journal Supplement Series, Vol. 148, p. 1-27, 2003.

Rubino-Martin, J. A. ; Chluba, J. ; Sunyaev, R. A. , Lines in the cosmic microwave background spectrum from the epoch of cosmological hydrogen recombination, Monthly Notices of the Royal Astronomical Society, Vol. 371, p. 1939-1952, 2006.

Chluba, J. ; Sunyaev, R. A. , Free-bound emission from cosmological hydrogen recombination, Astronomy and Astrophysics, Volume 458, p.L29-L32, 2006.

Staniszewski, Z. ; Ade, P. A. R. ; Aird, K. A. ; Benson, B. A. ; Bleem, L. E. ; Carlstrom, J. E. ; Chang, C. L. ;Cho, H. -M. ; Crawford, T. M. ; Crites, A. T. ; de Haan, T. ; Dobbs, M. A. ; Halverson, N. W. ; Holder, G. P. ;Holzapfel, W. L. ; Hrubes, J. D. ; Joy, M. ; Keisler, R. ; Lanting, T. M. ; Lee, A. T. ; Leitch, E. M. ; Loehr, A. ;Lueker, M. ; McMahon, J. J. ; Mehl, J. ; Meyer, S. S. ; Mohr, J. J. ; Montroy, T. E. ; Ngeow, C. -C. ; Padin, S. ;Plagge, T. ; Pryke, C. ; Reichardt, C. L. ; Ruhl, J. E. ; Schaffer, K. K. ; Shaw, L. ; Shirokoff, E. ; Spieler, H. G. ;Stalder, B. ; Stark, A. A. ; Vanderlinde, K. ; Vieira, J. D. ; Zahn, O. ; Zenteno, A. , Galaxy clusters discovered with a Sunyaev-Zel'dovich effect survey, eprint arXiv:0810.1578, 2008, Astrophysical Journal (in press)