CHAPTER 1

INTRODUCTION AND REVIEW OF LITERATURE

1.1 INTRODUCTION

Understanding cancer and cancer signaling has been a field of interest for researchers that have seen several advances with the development of potent tyrosine kinase inhibitors and also drawbacks especially with the failure of several drugs at clinical trials. Studies have been generally directed with a genomic perspective in mind with no investigation done in tandem to explore protein dynamics. Proteomics of signaling molecules done in keeping genomic perspectives is very necessary in understanding signaling kinetics. This work describes the investigation of the Epidermal Growth Factor Receptor (EGFR) signaling pathway which is an extensively studied signal transduction pathway in the context of mutant EGFR signaling conferring sensitivity to the EGFR inhibitor Erlotinib in Non Small Cell Lung Cancer (NSCLC). Lung cancer is the leading cause of mortality in India with cancers of the lung constituting 8.73% of relative proportion of all cancers in Bangalore (NCRP, 2008). The development of specific tyrosine kinase inhibitors (TKIs) such as Erlotinib against EGFR has been promising in targeted therapy of lung cancers. Specific driver mutations in the EGFR kinase domain are responsible for determining sensitivity or resistance to TKIs. More than 50% of the mutations have a leucine to arginine substitution at position 858 (L858R) that confers sensitivity to Erlotinib. EGFR activation at the membrane triggers a series of tyrosine and serine/threonine phosphorylation cascades. The goal of this study is to identify the sites of phosphorylation of proteins and quantify the degree of phosphorylation upon ligand stimulation and TKI inhibition of a lung adenocarcinoma cell line harboring a TKI-sensitizing mutant EGFR. A phosphoproteomic approach was followed to identify and quantitate change in phosphorylation kinetics and identify signaling molecules that are essential and play a vital role in signal transduction. This is the first report of such a detailed investigation from a proteomic point of view to understand signaling dynamics in the presence of a tyrosine kinase inhibitor Erlotinib and have the potential to identify
mechanisms of drug sensitivity and potential new targets for drug treatment. A review of literature is done in this chapter describing statistics and epidemiology of lung cancer progressing to histopathology and current treatment options with particular focus on Erlotinib with data from clinical trials followed by information on the EGF receptor and how mutations effect signaling in NSCLC was finally discussed followed by a current review on the use of mass spectrometry to study phosphorylation dynamics to decode signaling molecules that play an important role in the EGFR pathway.

1.2 LUNG CANCER: STATISTICS AND EPIDEMIOLOGY

Cancer is the leading cause of death in the world and accounted for nearly 13% of deaths in 2008 (7.6 million). Tobacco use is a major risk factor for cancer. Other risk factors include alcohol, poor diet and lack of physical inactivity. The main cancer however is of the lung which accounts for nearly 1.4 million deaths. Lung cancer is the second most common cancer among white, black, American Indian/Alaskan Native, and Asian/Pacific islander men and third among Hispanic men in the US according to the NCI cancer statistics (2010). Predicted statistical figures estimated a total of 1,529,560 new cancer cases and 569,490 deaths in the US in 2010. Overall cancer rates have however declined in both men (1.3% per year from 2000 to 2006) and women (0.5% per year from 1998 to 2006) due to a decrease in 3 cancer sites in men (lung, prostate, and colon and rectum/colorectum) and two major cancer sites in women (breast and colorectum). Cancers of the lung and bronchus have estimated to be 116,750 cases in men and 105,770 cases in women for the year 2010. Estimated deaths from lung bronchial cancer have been estimated to be at 86,220 and 71,080 cases in men and women respectively as shown in Figure 1. Cancers of the lung and bronchus are the most fatal of all cancers and is hence a very important problem to contend with in terms of chemotherapy and treatment.
1.3 Risk Factors Associated with Lung Cancer

There are several risk factors associated with lung cancer. The common risk factors associated with NSCLC are listed below

1. Cigarette smoking
2. Passive smoking
3. Diet and Food supplements
4. Alcohol
5. Physical activity
6. Pollution and
7. Predetermined genetic susceptibility

Tobacco smoking is the primary cause of lung cancers and will continue to be a major killer till global smoking levels reduce. In the US and UK alone there could be a reduction in lung cancer rates over the next two decades while passive smoking could account for nearly 1.6% of cancers of the lung. A recent study also shows that passive
smoking during childhood increases the risk of lung cancer by 3.6 fold. It has been shown in one study that of 399,767 patient and 3137 lung cancer cases, people who consumed alcohol at an average of 30g/day had a slightly higher risk for lung cancer. Alcohol consumption was associated with a greater risk in male never smokers as well. Available research suggests that physically active individuals have a lower risk of lung cancer. Moderate to high level of activity was associated to a 13-30% reduction in lung cancer risk. Thus, it is an important factor to reduce risk and mortality. Emissions rich in polycyclic aromatic hydrocarbons cause long term adverse effects on the lung possibly by inducing oxidative stress, inflammation and dysfunction of the autonomic regulatory system. The proportion of lung cancers arising due to air pollution was attributed to be 11% in Europe. Exposure to crystalline silica, chrysotile asbestos and other carcinogens increased the risk for lung cancer. Further uranium miners and nuclear plant workers had a higher risk as a result of exposure to particulate radioactive molecules while other familial clustering studies over 60 years suggested a hereditary link of lung cancer. Carriers of the P53 gene had an increased risk of cancer. Carriers who smoked cigarettes are three times more susceptible that those who did not smoke. A genome wide linkage study identified a major susceptibility locus for lung cancer in 6q23-25p. Other studies also showed a 30% higher risk for people with one copy of chromosome 15 marker (nicotinic acetylcholine receptor) and a higher risk of 70-80% for lung cancer for those with two copies.

1.4 LUNG CANCER HISTOPATHOLOGY

There are primarily 5 major classifications of lung cancer based on the histological presentation that are universally followed which are listed below:

1. Small Cell Lung Cancer
2. Large Cell Carcinoma
3. Adenocarcinoma and
4. Squamous Cell Carcinoma
5. Non Small Cell Lung Cancer
1.4.1 SMALL CELL LUNG CANCER

Small Cell lung cancer accounted for nearly 15% of all bronchiogenic carcinomas\(^{14}\). This form was primarily associated with various paraneoplastic syndromes, including syndromes associated with inappropriate anti-diuretic hormone secretion, paraneoplastic cerebellar degeneration and also with Lambert-Eaton Syndrome as well. As small cell lung cancers metastasized early the role of surgery in this case was very limited. The disease was highly sensitive to chemotherapy and rarely difficult to cure. Combination therapy involved platinum and etoposide in combination with thoracic radiotherapy. It was found to increase survivability by 5% over chemotherapy alone. In Japan and other Asian countries, combination treatment with Irinotecan and cisplatin is the standard chemotherapeutic regimen. Other clinical trials with imatinib and gefitinib have now shown promising results in clinical trials. Cells in this cancer presented a round to ovoid to spindle shape with not much cytoplasm evident. They have a high mitotic count and grow in clusters but do not show either glandular or squamous organization. Dense neurosecretory granules are seen under electron microscopy. These cells are immunoreactive to keratin, thyroid transcription factor 1 (TTF1) and epithelial membrane antigen. The most important cause of SCLC is smoking and accounts for nearly 95% of the cases\(^{15}\).

1.4.2 LARGE CELL CARCINOMA

Large cell neuroendocrine carcinoma is distinct from atypical and typical carcinoids and squamous cell carcinomas (SCC). Cell sizes are thrice as large as SCC cells and have an organoid growth pattern with cellular palisading or rosette like areas, have a very high mitotic rate and variable granular chromatin pattern\(^{16-17}\). Survival post resection is worse than other NSCLCs. Incidence of these cancers appeared to be between 2.1-3.5% of known lung cancers. Smoking is the primary cause of this cancer. Electron microscopy revealed that these cells have neurosecretory and occasional granular differentiation and intercellular junctions suggestive of squamous differentiation. The tumors stain for neuron specific enolase, carcinoembryonic antigen, keratin, chromogranin, LEU7,
synaptophysin and adrenocorticotrophic hormone. Post operative adjuvant therapy or radiotherapy has been used to treat these cancers but no definitive survival rates have been established for such therapies. Majority of lung neoplasms were shown to express the multidrug resistance gene (MDR1) in earlier studies\(^1\). Octrotide a somatostatin homolog was shown to control metastatic growth while being well tolerated in treatment of other neuroendocrine cancers\(^2\).

### 1.4.3 ADENOCARCINOMAS

Adenocarcinomas are the most prevalent form of lung cancer in young males and in women of all ages in never smokers and in former smokers. Differences in smoking habits appeared to favor development of distal bronchiolar and alveolar carcinogenesis at the expense of proximal SCC\(^{20-21}\). They are highly heterogeneous and only a few tumors show a pure histological pattern. They could be acinar, papillary, brochioalveolar or a solid carcinoma with mucin formation. Surgical resection is the primary mode of treatment. Tumors of several kinds are now included in the recent WHO/IASLC classification. Thyroid transcription Factor-1 (TTF-1) has been the primary marker for differential diagnosis for 85% of primary lung\(^2\).

### 1.4.4 SQUAMOUS CELL CARCINOMA

SCC is characterized by its histopathological resemblance to the epidermis. The cells are large, flattened stratified cells that have abundant cytoplasm. There is formation of intra-cytoplasmic keratin which caused the formation of keratin pearls and intercellular bridges/prickles. Most SCC arises centrally in the lungs in the main lobar, segmental or sub-segmental bronchi. The carcinomas were usually firm off white masses in the lumen of the bronchus. They invade the underlying wall and parenchyma. There is a central softening of the cavitations of the tumor due to necrosis or keratinous debris. Histochemical stains for keratin, epithelial membrane antigen, carcinoembryonic antigen and involucrin are positive. Rare variants include spindle SCC, papillary variant tumors. Small cell squamous carcinomas are possibly a variant of SCC\(^2\).
1.4.5 Non Small Cell Lung Cancer (NSCLC)

NSCLC accounted for nearly 85% of all lung cancer cases. Lung cancers have become more frequent among former than current smokers. Although cigarette smoking has decreased in the US, it has dramatically increased in areas such as China and other developing countries where it has reached high levels of usage\textsuperscript{24}.

1.4.5.1 Staging

Staging is an important criterion to determine appropriate therapy for NSCLC. Patients with Stage I or II benefit from resection but patients with an advanced stage of the disease require other modes of non-surgical treatment. A CT scan of the thorax and upper abdomen can be done primarily but have limited sensitivity in cases of microscopic metastatic disease and cannot distinguish mediastinal lymph nodes that are enlarged due to benign reactive hyperplasia. A PET scan with fluorine labeled fluorodeoxyglucose has a greater sensitivity to detect metabolically active malignant disease and can lead to changes in initial changes and treatment of NSCLC\textsuperscript{25-27}. Mediastinoscopy or thoracotomy is considered a standard criterion for mediastinal staging of lung cancer. Newer preoperative staging tools like endoscopic bronchial ultrasonography and esophageal ultrasonography for guided biopsies are also done and are less invasive to mediastinoscopy\textsuperscript{28}. Cervical Mediastinoscopy is the current standard in preoperative nodal staging of lung cancer with nearly 93% sensitivity\textsuperscript{29}.

1.4.5.2 Screening

Lung cancer has a 5-year survival rate of 15% and hence timely detection and accurate staging is very necessary to determine treatment options. There are currently no guidelines for mass screening by radiographic techniques as the American Cancer Society does not recommend screening for at-risk individuals. Large scale studies are underway to determine the benefits of mass screening.
1.4.5.3 TREATMENT

There are several recommended modes of treatment for NSCLC and the most important of them are primarily surgery, chemotherapy, radiotherapy and targeted drug therapy. These are discussed in brief in this section.

1.4.5.4 SURGERY

Surgery is the most consistent and successful option for cure in case of patients diagnosed with lung cancer. Completely resectable tumors favor this option and it is important that the patient must be able to tolerate the surgery. Success of surgery is dependent on complete resection of the tumor. Studies have indicated the benefits of patients undergoing lobectomy to over those who prefer limited resections (wedge resection or segmentectomy) with an increased survival rate in the former case. Lobectomy remains the standard mode of resection of NSCLC. Minimal access surgical procedures are also expanding applicability of resection.

1.4.5.5 CHEMOTHERAPY

Nearly 70% of patients presented with advanced or metastatic stages of the disease at time of diagnosis and adjuvant chemotherapy were generally recommended for patients with resected stages IIA through IIIA NSCLC. Patients with resected lung cancer had very high chance of relapse and a 5-year survival benefit was shown in patients receiving platinum-based chemotherapy and has initiated several adjuvant trials. Platinum-based chemotherapy is considered as the first line of defense for NSCLC cancer patients and has been associated with a moderate improvement in the quality and survival of life. However, patients soon develop resistance to the treatment resulting in limited overall survival in spite of effective regimens and a second-line treatment options. There have been several lung cancer trials that have been initiated with Tyrosine Kinase Inhibitors (TKIs) as primary line of defense or in combination with platinum-based doublets or with TKIs alone for maintenance chemotherapy.
1.4.5.6 ADJUVANT CHEMOTHERAPY

Patients with resected lung cancer have very high chance of relapse. A 5-year survival benefit was shown in patients receiving platinum based chemotherapy and has initiated several adjuvant trials as shown in Table 1. Platinum based chemotherapy is considered as first line of defense for and has been associated with a moderate improvement in the quality and survival of life. However, patients soon develop resistance to the treatment resulting in limited overall survival in spite of effective regimens and second-line treatment options. There have been several lung cancer trials that have been initiated with TKIs as the primary line of defense or in combination with platinum-based doublets or with TKIs for maintenance chemotherapy as shown in Table 2.
<table>
<thead>
<tr>
<th></th>
<th>Adjuvant Combination</th>
<th>Stage</th>
<th>5-year survival</th>
<th>Patients</th>
<th>Trial</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjuvant Chemotherapy Group</td>
<td>Observation Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>CDDP + ETP or + VNR or + VBL</td>
<td>I-III</td>
<td>45</td>
<td>40</td>
<td>1867</td>
<td>IALT</td>
</tr>
<tr>
<td>2</td>
<td>CDDP + VNR</td>
<td>IB-II</td>
<td>69</td>
<td>54</td>
<td>482</td>
<td>NCIC JBR10</td>
</tr>
<tr>
<td>5</td>
<td>CDDP + CPA + VCR + DOX + UFT</td>
<td>I-III</td>
<td>49</td>
<td>31</td>
<td>70</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>CDDP + VDS + UFT</td>
<td>II-III</td>
<td>38</td>
<td>37</td>
<td>95</td>
<td>OLCS</td>
</tr>
<tr>
<td>8</td>
<td>CDDP + VDS + UFT</td>
<td>I</td>
<td>88</td>
<td>66</td>
<td>150</td>
<td>ACLC</td>
</tr>
</tbody>
</table>

Table 1 Clinical trials that have been conducted till date for cancers of the lung. (CDDP-Cisplatin, ETP-Etoposide, VNR- Vinoreline, VBL- Vinblastine, PTX- Paclitaxel, CPA-Cyproterone acetate, DOX- Doxorubicin, UFT- Tegafur uracil and VDS-Vindesine)
<table>
<thead>
<tr>
<th>Trial Name</th>
<th>Cohort</th>
<th>Treatment</th>
<th>Result for TKIs</th>
<th>Reference</th>
</tr>
</thead>
</table>
| Single agent EGFR TKI first line chemotherapy for NSCLC
1 | IPASS | Adenocarcinoma; Asian non smokers | Gefitinib (n=609) Vs Carboplatin/paclitaxel (n=608) | Positive | Mok, T.S., et al (2009)\(^{31}\) |
2 | WJTOG3405 | Japanese; EGFT mut positive | Gefitinib (n=86) Vs Cisplatin/Docitaxel (n=86) | Positive | Mitsudomi, T., et al (2010)\(^{42}\) |
3 | NEJ002 | Japanese; EGFT mut positive | Gefitinib (n=98) Vs Carboplatin/paclitaxel (n=96) | Positive | Maemondo, M., et al. (2010)\(^{43}\) |

| EGFR-TKI combined with platinum-based doublets in the first-line chemotherapy for NSCLC
1 | INTACT I | None | Gefitinib with Cisplatin/Gemcitabine (N= 365 for Gefitinib 500 mg/day; N= 365 for 250mg/day) Vs Cisplatin/Gemcitabine (N= 363) | Negative | Giaccone, G., et al. (2004)\(^{44}\) |
2 | INTACT II | None | Gefitinib with Carboplatin/Paclitaxel (N= 347 for Gefitinib 500 mg/day; N= 345 for 250mg/day) Vs Carboplatin/Paclitaxel (N= 345) | Negative | Herbst, R.S., et al (2004)\(^{45}\) |
3 | TRIBUTE | None | Erlotinib with Carboplatin/Paclitaxel (N= 539) Vs Carboplatin/Paclitaxel (N= 540) | Negative | Herbst, R.S., et al. (2005)\(^{46}\) |
4 | TALENT | None | Erlotinib with Cisplatin/Gemcitabine (N= 580) Vs Cisplatin/Gemcitabine (N= 579) | Negative | Gatzemeier, U., et al. (2007)\(^{47}\) |

| EGFR-TKI in the maintenance chemotherapy for NSCLC
1 | SWOG0023 | None | Gefitinib (N= 118) Vs (after Cisplatin/Etoposide/RT followed by Docetaxel) Placebo (N=125) | Negative | Kelly, K., et al. (2008)\(^{48}\) |
2 | WTJOG0203 | Japanese | Gefitinib (N= 298) Vs (after Platinum-based doublets) Placebo (N=297) | Positive in Adenocarcinoma | Takeda, K., et al. (2010)\(^{49}\) |
3 | SATURN | None | Erlotinib (N= 438) Vs (after Platinum-based doublets) Placebo (N=451) | Positive | Cappuzzo, F., et al. (2010)\(^{50}\) |
4 | ATLAS | None | Erlotinib/Bevacizumab (N=370) Vs (after Platinum-based doublets/Bevacizumab) Bevacizumab (N= 373) | Positive | Miller VA, et al. (2009)\(^{51}\) |

Table 2 Clinical trials till date for cancers of the lung with all known studies on adjuvant therapy
1.4.5.7 Radiotherapy

With a high risk of systemic spread, local therapy either in the form of surgery or radiotherapy can be curative for certain forms of lung cancer. The cure local regional failure rates (LRFs) for N0 disease ranges from 6-28%, 18-49% for N1 disease, and 6-65% for N2 disease\(^5\). Post operative local/regional radiotherapy (PORT) has been shown to be effective in combating NSCLC. Risks are higher in case of total gross resection. PORT is recommended for patients with pathologic N2 disease or those with close/positive margins or residual macroscopic disease. For N0 patients, it is considered as a therapy of choice. However RT must be applied carefully to minimize risk\(^5\).

1.4.5.8 Targeted Therapy

Several signaling molecules have been shown to be differentially regulated in NSCLC. Among them current therapies target the Vascular Endothelial Growth Factor (VEGF) and the Epidermal Growth Factor Receptor (EGFR). VEGF binds to the VEGF Receptors VEGFR1 and VEGFR2 and activation of VEGFR2 is alone sufficient to induce VEGF based mitogenesis, angiogenesis and vascular permeability\(^3\). Bevacizumab (Avastin, Genentech) is known to be effective in combination with Paclitaxel and Carboplatin in people who have not yet received chemotherapy\(^4\). Among the EGFR inhibitors Gefitinib (Iressa, AstraZeneca) and Erlotinib (Tarceva, Genentech) have been shown to be effective drugs.

1.5 ERBB Receptor Super Family of Receptors

EGFR belongs to the ERBB family of Receptor Tyrosine Kinases (RTKs). These receptors are only found in metazoans in stark contrast to serine threonine kinases that are conserved in eukaryotic and also in unicellular and multicellular organisms\(^5\). EGFR was the prototypical receptor tyrosine kinase to be first cloned. It belongs to a large family of transmembrane receptors where for the first time, the importance of ligand mediated oligomerization on the activation on an enzyme was shown. The ERBB family originally named because of their homology to the erythroblastoma viral gene product, v-ERBB consists of four members collectively called as the HER or ERBB family. EGFR,
HER2/Neu, ERBB3/HER3, ERBB4/HER4. As with all RTKs, the domain structure primarily consists of an extracellular region, a single spanning transmembrane domain and intracellular juxtamembrane region, a tyrosine kinase domain and a C-terminal regulatory region. There are two main groups of ligands that regulated the ERBB receptors. The contain the EGF agonists that activate EGFR and the neuregulins that bind and regulate ERBB3 and ERBB4\(^{56}\). There are at least seven known EGF agonists that include EGF, transforming growth factor alpha, amphiregulin, betacellulin epigen, epiregulin and Heparin binding EGF-like growth factor. Some of these ligands are bi-specific and can activate ERBB4, ERBB3 and ERBB4. ERBB2 however has not shown to have a soluble ligand and is hypothesized to form heterodimers with other ERBB members\(^{56-57}\) (Figure 2). The activation/deactivation of RTKs determines the modular control of growth, differentiation, cell motility and or survival of a cell\(^{58}\). Ligand induced dimerization activates autophosphorylation by stimulated intracellular kinase domains\(^{59-61}\). Post phosphorylation of tyrosine residues on the protein moiety assist in serving as docking sites for several SH2/PTB domain containing proteins that transmit signals to downstream proteins\(^{56,62-64}\). The growth factor receptor bound 2 (GRB2) and Src

![Figure 2: ERBB family of receptors along with their respective agonists. Rowinsky EK Annual Reviews (2004) 55: 433-457](Image)
homology 2 containing (SHC) are responsible for the recruitment of RAS and activate the MAP Kinase cascades.

1.6 EGFR INHIBITORS CURRENTLY IN USE

Among the most potent inhibitors Gefitinib and Erlotinib are two of the most successful inhibitors that are currently in use.

1.6.1 GEFIGINIB

Gefitinib (Iressa, AstraZeneca) was first introduced in May 2003 with FDA approval. It was used for therapy after failure of chemotherapy in patients with advanced or metastatic NSCLC. Gefitinib is an anilinoquinazoline (4-quinazolinamine, N-(3-chloro-4-fluorophenylamino)-7-methoxy-6-[3-(4-morpholinyl) propoxy]). Although it acts on EGFR, the mode of action has not been well characterized as of today. Pre-incubation of A431 cell extract with gefitinib prior to EGF treatment inhibited autophosphorylation of EGFR TK in a dose-dependent manner and the behavior is also observed in human tumor cell lines from the lung, prostate, colon, and head and neck cancers. These results correlated with inhibition of tumor xenograft growth in vivo.

1.6.2 ERLOTINIB

Erlotinib (Pfizer and OSI Pharmaceuticals) is also an orally available reversible, ATP-competitive inhibitor of the EGFR tyrosine kinase (TK) whose chemical name is ([6, 7-bis (2-methoxy-ethoxy)-quinazolin-4-y1]-[3-ethylphenyl] amine) (Figure 3). It is similar in function to gefitinib as an EGFR Inhibitor. It has also been shown to delay tumor growth in human tumor xenografts. This quinazoline derivative binds competitively to the ATP site in the EGFR TK domain. The IC₅₀ for EGFR ranges from 2-20 nMol/l with purified kinase or in cell culture. Other TKs were blocked with concentrations that were 1000 times higher. In preclinical models, exposure to the drug causes upregulation of p27, G₁/G₀ arrest and induction of apoptosis. A dose dependent inhibition of growth was seen in the HN5 cell line which is an EGFR rich HNC model. Inhibitory growth was seen in concentrations of 12.5 mg/kg/day for 20 days. Maximum inhibitory effect on EGFR phosphorylation was turned down 80% in 1 hour post treatment with the drug and
inhibition was seen at levels of 70-80% for approximately 12 hours and resumed baseline levels in 24 hours\textsuperscript{67}.

\textbf{Figure 3:} Chemical structure of Erlotinib

The ED\textsubscript{50} for the drug is 10mg/kg and there were no different in routes of administration be they oral or peritoneal in mice. Plasma concentrations in mice were 2.9-100pmol/L at doses of 2.9-9.2 mg/kg. A linear relationship is observed between target inhibition and \textit{in vivo} studies in mice.

\textbf{1.6.2.1 PHASE I AND II CLINICAL STUDIES OF ERLOTINIB}

Erlotinib has dose-dependent pharmacokinetics with daily dosing not leading to drug accumulation. A dose of 150 mg/day was determined to be the maximum-tolerated dose at which biologically relevant plasma levels were achieved. A phase II trial was conducted to evaluate Erlotinib in advanced refractory NSCLC \textsuperscript{68-69}. Results showed complete responses in two patients (4%), partial responses (PRs) in five patients (9%), and prolonged stable disease (SD) in 22 patients (39\%). The median survival time (MST) was 8.4 months. Other associated lung cancer symptoms (fatigue, dyspnea, and cough) improved with the use of Erlotinib. In advanced NSCLC patients, Erlotinib as a single agent has also been tested as first-line treatment\textsuperscript{70}. Fifty-three chemotherapy-naive patients with stage IIIB/IV NSCLC received oral Erlotinib (150 mg/day). The overall rate of nonprogression at 6 weeks was 52.8\% (28 of 53 patients). The objective response (OR) rate was 22.7\%, and the MST was 391 days. Women were more responsive than men and occurred mostly in those with adenocarcinoma (including bronchioloalveolar carcinoma histology) and in nonsmokers. However, responses were also observed in patients with other tumor histology and in former or current smokers. A phase II study was conducted to evaluate Erlotinib as first-line monotherapy in 80 chemotherapy-naive greater than 70
years of age patients with stage III/IV disease\textsuperscript{71}. There were eight PRs (10%), and 33 patients (41%) experienced SD for 2 months or longer. The MST was 10.9 months. The 1- and 2-year survival rates were 46% and 19%, respectively \textsuperscript{72}.

1.6.2.2 PHASE III CLINICAL STUDIES OF ERLOTINIB

Chemotherapy offers symptomatic relief and a modest improvement in survival. Few options are available for the treatment of patients with progressive disease after failure of second-line therapies. The clinical activity of Erlotinib chemotherapy-refractory NSCLC patients effectively prolonged survival. The National Cancer Institute of Canada conducted a phase III randomized trial, BR.21, in which Erlotinib was compared with placebo in stage III/IV NSCLC patients who had failed first- or second-line chemotherapy. A total of 731 patients were randomized in a 2:1 ratio to receive either Erlotinib at 150 mg/day or placebo. Those patients had metastatic NSCLC that had previously been treated with one standard chemotherapy regimen (50% of patients) or with two chemotherapy regimens (50% of patients). All patients received platinum-based chemotherapy. The OR rate was 8.9% in the Erlotinib arm and <1% in the placebo group (p < .001). The median durations of response were 7.9 months and 3.7 months, respectively. The MST was 6.7 months for those in the Erlotinib regimen compared with 4.7 months for those in the placebo arm (p < .001; hazard ratio [HR], 0.7; 95% confidence interval [CI], 0.58–0.85). ORs were more frequent in women (14% versus 6%; p = .0065), in patients with adenocarcinoma, as compared with other histotypes (14% versus 4.1%; p <0.0001), and in patients without a smoking history (25% versus 4%; p < .0001)\textsuperscript{73-74}.

1.7 MUTATIONS IDENTIFIED IN THE EGF RECEPTOR

EGFR has been shown to be overexpressed in 40-80\% of NSCLC tumors and other epithelial tumors\textsuperscript{75}. Gefitinib and Erlotinib target the ATP cleft and competitively inhibit binding of ATP to the ATP cleft of EGFR TK. Auto and transphosphorylation of the receptor leads to the recruitment of downstream effectors and helps with the activation and proliferation of cell survival signals. Though the molecule is ubiquitously expressed, inactivation of the gene cause minimal defects which is an added advantage for the use of
targeted inhibitor therapies\textsuperscript{76}. Initial studies with gefitinib have shown in NSCLC patients who showed positive responses to the drug had mutations in the EGFR gene indicative of the receptor pathway to be important in this particular cancer. In frame deletions were reported in the initial studies of Exon 19 (746-750, 747-751 and 747-753) and amino acid substitutions at the 858 residue from a leucine to an Arginine residue and a L861Q mutation. A fourth missense mutation was also observed that resulted in the substitution of a glycine with a cysteine residue. The matched normal tissue from the same patients showed only WT EGFR. Exons 19 and 21 were known to be hotspots that harbored these mutations. In vitro models with the mutant receptors expressed in Cos-7 cells showed that on EGF treatment, the activity of the kinase doubled or tripled with respect to the WT receptor. These mutations did not affect the stability of the protein though. The quantitation was done by monitoring the Y1068 residue and EGFR antibody based assays. In comparison with the WT EGFR expressing cells where the EGFR phosphorylation levels were attuned to the normal levels 15 minutes post activation in the mutant receptor expressing cells, the activation continued for nearly three hours post activation\textsuperscript{77}. Seven of the eight mutations reside in the ATP cleft and it was hypothesized that these mutations governed the sensitivity to gefitinib.

Other studies with lung adenocarcinoma and bronchioloalveolar cell lines were done to determine sensitivity of these cells lines to gefitinib. Studies with a sensitive cell line H3255 which was derived from the malignant pleural effusion of a Caucasian female non smoker. The cell line was more sensitive than other cells that were investigated with an IC\textsubscript{50} of 40nM in a 72 hour survival assay. Treatment with 100mM gefitinib completely inhibited phosphorylation of EGFR in H3255. Downstream targets such as extracellular signal regulated kinase 1/2 (ERK 1/2) and also the v-akt murine thymoma viral oncogene homolog (AKT) were inhibited. The reason for such a behavior was attributed to the L858R mutation seen in H3255\textsuperscript{78}.

Another study at the same time documented the mutations also being the reason for sensitivity in Erlotinib treated cancers. Additional mutations were also seen which were not reported in gefitinib treated cancers. Deletion from 747-752, double mutations were also seen in one sample which was a C to T change substituting cysteine for Arginine an
R776C change. These mutations were common in non smokers who had lung adenocarcinomas. Erlotinib treatment of the L858R mutant expressing 293 cells showed a dose dependent inhibition of receptor activity.

These mutations are thus important in determining sensitivity to TKIs like gefitinib and Erlotinib. The study showed none of these mutations were observed in patients of an East Asian origin. A summary of all mutations from these studies have been summarized below in Figure 4.

1.7.1 MUTATION STATUS IN THE ASIAN POPULATION

There is currently only one report on the study of these mutation in the Indian population. This study had 220 patients who were enrolled in this study from January 2008 to July 2010. FFPE samples were obtained and RT-PCR was done to ascertain their mutational status. Highest mutation rate was noted for Exon 19 deletions which accounted for 51.6 of the cases followed by Exon 21 L858R mutations that accounted for 26.2% of cases. Though the T790M mutation was not analyzed, its prevalence was 3% of all cases. Other rare mutations were also seen like the S768I in exon 20 (5.5%), L861Q in exon 21(1.6%).

---

**Figure 4**: Summary of all known mutations that confer resistance or sensitivity in lung adenocarcinomas

1.8 HIGH THROUGHPUT STRATEGIES FOR PROFILING OF CANCER

The post-genomic era ushered in hopes of discoveries in the fields of medicine and allied sciences for curing disease and rectifying genetic defects. Identification of minor errors in the genomes biological instructions either inherited or acquired during one’s life can make all the difference between health and disease. Biomarker discovery using clinical genomics and proteomics approaches will possibly enable us to detect these errors in people and identify those at highest risk of an ailment and allow them to reduce that risk by making lifestyle changes or taking preventive drugs. Advancement in new technologies in the field of clinical genomics has opened up new avenues for identifying biomarkers which can be used for disease risk assessment, early detection, prognosis and for assessing drug response and toxicity. Biomarkers are expected to be highly accurate, efficient, and reliable for assessing disease risk. Hence the requirement today is rapid, efficient, and systematic approaches for searching biomarkers that are potential candidates with high accuracy for disease diagnosis/prognosis. Scientists are getting new insight and ideas about how to design more effective drugs by studying how these errors actually cause disease. High throughput molecular profiling of cancer and cancer pathways has been indicated as the next step in identification of biomarkers in the fight against cancer post the human genome project which would enable scientists to investigate gene and gene products that determine or underlie tumor development and progression53-55.

Development of high throughput methods of sample preparation and automated sample processing facilitate correlating genomic with proteomic information80-81. Development of Laser Capture Microdissection (LCM), protein arrays, antibody arrays and small molecule arrays have had a substantial effect on proteomic profiling of cancers82-85. Gene expression studies provide information on responses of cells to cancer however, more validation is needed in experimental models and further clinical trials need to be done before routine clinical use. No information can be obtained from these experiments on dynamic processes in the cell such as phosphorylation, protein trafficking, and protein-protein interactions in cancer cells post therapeutic stress86.
High throughput protein arrays provide this information on differentially expressed proteins when the normal is compared with the cancer tissue as in the breast, AML, and other diseases. Development of profiles on cancers using mass spectrometry has been suggested earlier and ICAT (Isotope Coded Amino Affinity Tags) based MS studies have been done to study ovarian cancer, and primarily for comparative proteomics using either cell lines or tissues, study metastatic progression and metabolite profiling.

1.9 SIGNALING AND BIOMARKERS

Research in our lab involves in phosphotyrosine profiling and comparison of levels of tyrosine phosphorylated proteins in the context of signaling mechanism in lung cancer, gastrointestinal stromal tumor and pancreatic cancer using high throughput mass spectrometry approach. Initially, these target proteins were identified by immunoaffinity purification using phosphotyrosine specific antibodies (4G10, PY100, and PY20). Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) based quantitative proteomics strategy was employed for the dynamic changes in specific proteins up to five different states including control, ligand stimulation and drug inhibition. The use of quadrupole time-of-flight (qTOF) and LTQ Orbitrap XL ETD mass spectrometer for these studies was essential in getting high quality data. Since thousands of such peptides can be quantified by LC-MS/MS, to date this is most effective way to study phospho signaling. Our lab has been involved in efforts to identify specific sites governing cell proliferation and sensitivity to drug treatment.

A major standardization that has been done in-house is enrichment of phosphotyrosine containing peptides using specific antibody (pY100, CST; Phosphoscan) followed by quantitative proteomics analysis using mass spectrometry. Using this approach we have been able to identify as well as quantify hundreds of phosphotyrosine peptides including previously not reported sites in lung cancer. Further, these studies have been complemented by performing a global analysis of phosphorylation sites using metal affinity chromatography based enrichment all phosphopeptides, which mainly includes titanium dioxide (TiO2) and Immobilized Metal Affinity Chromatography.
(IMAC) methods for enrichment of phosphopeptides. In addition, protein fractionation using reversed phase LC, strong cation exchange chromatography for fractionation of peptides are routinely used. Identification of phosphoserine and phosphothreonine sites are still challenging when there more than one potential sites in peptide. Recently, our lab has used electron transfer dissociation (ETD) in contrast to usual collision induced dissociation (CID) \(^{111-114}\) method of fragmentation in ion trap mass spectrometer to identify such sites.

**1.10 MASS SPECTROMETRY AND LUNG CANCER**

Clinical proteomics involves reproducible analysis of human samples thus requiring a large sample repository to start with. Initial studies in the human plasma of lung adenocarcinoma patients involved in 2D Micro LC/MS analysis of the human plasma on a linear ion trap mass spectrometer (ITMS) using samples from health and diseased patients. The study revealed several group specific proteins that could be candidate markers specific to lung adenocarcinoma\(^ {115}\). Plasma biomarkers have been detected using monoclonal antibodies in NSCLC patients serum using ELISA\(^ {116}\), in detecting an EGFR signature in mouse models \(^ {117}\). Using label free methods, phosphoproteomics profiling identified onco Ras and its involvement in lung adenocarcinoma \(^ {118}\). Earlier studies used MALDI-MS of preinvasive lesions for detection of lung cancer for monitoring high risk individuals for surveillance and chemoprevention trials \(^ {119}\). Protein tyrosine kinases and phosphatases play a major role in cell signaling and the development of inhibitors to these kinases or phosphatases have enabled the need to study global tyrosine phosphorylation events that would likely lead to the identification of novel targets in drug discovery or target validation \(^ {120-121}\). 2 D peptide mass fingerprinting were some of the less sensitive analyses were used to identify biomarkers in cancerous and adjacent tumor tissues \(^ {122}\). Subcellular analyses were used to compare protein profiles in lung adenocarcinoma cells and in human bronchial epithelial (HBE) cells followed by pathway analysis and biological validation to reveal the Epithelial-Mesenchymal transition (EMT) phenotype shift in A549 cells with respect to HBE cells.
1.11 Mass Spectrometric Investigation of Phosphorylation Kinetics

Mass spectrometric techniques helped bridge the gap between biochemical and molecular aspects of cell signaling pathways. Traditionally, Edman sequencing has been used to identify protein molecules post biochemical purification. The bioavailability of the molecule is the limiting factor as the technique requires large amounts of the purified protein. This cannot always be the case in case we need to study or investigate substrates that are present in very low levels in the cells. Mass spectrometry has been a very important tool in helping overcome this hurdle by being able to identify and also quantitate proteins at attomole or femtomole levels\textsuperscript{58,123-124}. Mass spectrometry has been suggested as a definitive method to study multiprotein complexes obtained post affinity based purification. There are several mass spectrometric methods that are commonly used which are quantitative in nature. Among them SILAC\textsuperscript{97,125-127}, Isotope Coded Affinity Tags (ICAT)\textsuperscript{128}, Isobaric Tag for Relative and Absolute Quantitation (iTRAQ), and $^{18}$O labeling are the most common strategies used. Sample preparation is the key to any proteomic experiment dealing with phosphorylation kinetics requiring a higher diligence of sample preparation. There have been specific enrichment protocols that employ antibodies for immunoprecipitation of phosphorylated proteins/peptides or the use of specific IMAC, TiO$_2$ or graphite powder. There has also been the use of several labeled and non labeled quantitative methods that have been employed for protein quantitation and SILAC for quantitative phosphorylation monitoring are popular.

SILAC is a non hazardous method wherein the complete proteome is labeled in cells growing in cell culture. Mammalian cell lines are grown in a media lacking essential amino acids that are supplemented with a non-radioactive isotopically labeled for the same amino acid. The cells are not physiologically affected by the substituted amino acid. There is no visible change in cell morphology, doubling time or even the ability to differentiate\textsuperscript{97}. Furthermore, being more versatile than microarrays, it has been extensively used in studying proteins in cancers because of its advantage in being coupled with mass spectrometry to investigate several hundreds of proteins from a single
experiment. Expression levels for more than 440 proteins were investigated in microsomal fractions of prostate cancer cells with different metastatic potentials\textsuperscript{129}.

It is a good system to study protein modifications and was first used to study methylation patterns in cells and identified several sites not identified earlier increasing the list of methylation sites known in cells\textsuperscript{130}. Protein phosphorylation has been extensively studied as it is easy to compare complete proteomes of differentially labeled cells in a single experiment\textsuperscript{131}. A phosphoproteomic analysis of HeLa cells was first done to identify potential kinase substrates without having prior knowledge of signaling pathways and can was used for the targeted study of kinases or protein molecules in a general fashion\textsuperscript{93-94}. Coupled with RNAi, it was used to study temporal changes in proteins and specific substrates of proteins in a particular pathway as well\textsuperscript{132-133}. This system of labeling cells is not limited to mammalian cells but has been used also for quantitating proteins in yeast identifying nearly 2000 proteins\textsuperscript{134-135}. Whole organisms have also been labeled with SILAC amino acids incorporated into their diets in mice\textsuperscript{136} and flies\textsuperscript{137}. Further advancements in SILAC have been used to quantitate proteins with accuracy to the attomole level (Absolute SILAC)\textsuperscript{138} and for studying protein turnover dynamically with dynamic and pulsed SILAC\textsuperscript{139-140}.

A SILAC experiment consists primarily of three important steps.

- a. Adapting cells to heavy SILAC media
- b. Biochemical processing or enrichment or treatment of cells
- c. Mass spectrometry and data analysis.
Figure 5: Embryonic kidney 293T cells were grown in heavy lysine ($^{13}$C$_6$) containing medium after initial growth in normal medium on day 0. Aliquots of cells were removed every 24h, lysed, run on a gel, trypsin digested and analyzed using mass spectrometry. Doubly charged ions at m/z 760.4 and 763.4 correspond to a peptide from the protein adenylate kinase 2. The observed mass difference is because of incorporation of heavy Lys. As seen, by day 4, there is near complete incorporation of heavy Lys. Harsha H.C et al., 2008 Nat Prot 3 (3) 505-516
Adapting cells to SILAC media is the critical step for the experiment. Complete incorporation is very necessary for any experiment as incomplete incorporation can lead to biases in the experimental system by giving wrong quantitative values for proteins skewing results and making it difficult to interpret the results. **Figure 5** shows the complete incorporation of heavy lysine in adenylate kinase 2 protein. It is only after checking for complete incorporation is any cell treated or used for the experiment. Cell processing is the next crucial step and depends on the experiment. Cells can be treated with specific drugs and or growth factors and can be lysed in specific buffers. Antibody, protein based or chemical enrichment can be done to isolate the specific subpopulation of proteins of interest and also sub cellular fractionations can be done at this stage so as to isolated organelle fractions that can then be processed downstream depending on the experiment. Further, the lysates can be digested cleaned and run on the mass spectrometer.

### 1.12 Phosphorylation and the EGF Receptor

The low concentration of phosphoproteins in biological samples, inappropriate sample preparation, transiency of a phosphorylation biochemical reaction, inefficient automatic mass precursor ion selection or insufficient data handling are some of the key issues that need to be taken into consideration while investigation phosphorylation dynamics. Reversible phosphorylation is a highly regulated event in many tumorogenic cells. Phosphorylation of the receptor takes place in minutes and no major total protein changes occur at this time point and long term changes are only visible in long term experiments. The EGF receptor is located on the on the cell membrane and on interaction with the appropriate ligand i.e. EGF or TGF alpha, triggers dimerization, reversible phosphorylation, overexpression and translocation of the receptor molecules. However, there is a fine transient interplay between phosphorylation and dephosphorylation events that involve a fine play between phosphatases and the phosphorylated substrates that they bind with. This dynamics is a major issue that needs to be taken into consideration while designing sample prep and enrichment protocols. A comparative analysis of the phosphorylated and the non phosphorylated
peptide is essential to identify true peptides that are stimulated on EGF treatment for an unbiased comparative study taken that similar ionization efficiencies assumed. The stochiometric levels of EGFR must cross the MS detection limit and it is taken that a reasonable level would be 30 fmol EGFR prior to analysis. Studies in hepatocytes showed an average of 5000-10000 copies per cell are satisfactory for MS\textsuperscript{141} and this is a major limitation in biological samples that involve human tissues. Thus there is a need to introduce several fractionation and enrichment steps in combination with strong cation exchange chromatography (SCX) and metal affinity purification procedures such as IMAC and TiO\textsubscript{2} protocols. Fractionation at the protein level and the peptide level implements the need and increases phospho peptide detection levels. The primary study that was done in this context was using a pull down experiment employing the SH2 domain of GRB2\textsuperscript{150}. This experiment was done with SILAC labeled cells that were left untreated or treated with EGF and the lysates were subjected to a protein IP. A total of 228 proteins were identified of which, 28 were activated on EGF treatment. Further peptide IP experiments were used to identify substrates of ERBB receptors\textsuperscript{151}. STAT5 was found to be direct interactors with EGFR. Antibody immunoprecipitations of EGFR was used to understand interacting substrates of the receptor in several studies\textsuperscript{152-154}. Further in complementation with a protein phosphatase PTP\textsubscript{1B} the effect on the EGF and PDGF pathway was studied\textsuperscript{155}. The largest study to date using SILAC cells and 5 time points of EGF stimulation was also done to identify nearly 31 novel effectors. This was a comprehensive study of EGF determined phosphorylation that indicated receptor autophosphorylation within seconds of induction and SHC tyrosine phosphorylation occurring next indicative of temporal clones for interaction\textsuperscript{156}. A study by Zhang et al employed both iTRAQ labeled states that were subjected to phosphopeptide enrichment followed by IMAC and a q-TOF based analysis. Data was organized into clusters of phosphorylation sites using Self Organizing maps (SOMs) that recapitulated earlier biological findings and could be extended to find other proteins by means of extension and identify several new targets\textsuperscript{157}. In particular reference to NSCLC, a global survey of cancer cell lines identified nearly 4500 phosphotyrosine sites on around 2700 proteins. RTKs like PDGFR, ALK, ROS and DDR\textsubscript{1} were identified in this study\textsuperscript{158-159}. In combination with SCX fractionation and TiO\textsubscript{2} enrichment identified nearly 6600 sites on
2244 proteins with more than a 1000 sites that were activated two fold on EGF treatment\textsuperscript{144}.

### 1.13 Signaling Pathways in Adenocarcinoma Cell H3255

The H3255 cell line was one among the few cell lines that were generated from patients with NSCLC who were treated with 96 hour infusions of paclitaxel followed by bolus cisplatin\textsuperscript{160}. The cell line is resistant to concentrations of paclitaxel and docetaxel greater than 10\(\mu\)M. Elevated levels of Her2/Neu were observed by immunochemical studies. MDR1(Multi Drug Resistance) mRNA was also identified in the cell line. The cell line is sensitive to gefitinib at a concentration of 40nM and harbors the L858R mutation on EGFR. Further, there is amplification of EGFR with a 11 fold overexpression. There are no significant copy number changes suggesting that the L858R allele is preferentially amplified in H3255. It is possible that EGFR amplification with the mutation contributes to sensitivity to gefitinib. EGFR is constitutively phosphorylated and on EGF treatment is further hyperphosphorylated and dephosphorylation was observed on treatment with gefitinib. AKT is constitutively phosphorylated and is dephosphorylated on gefitinib treatment. Further, ERK 1/2 is not constitutively phosphorylated and is hyperactivated only on EGF treatment and completely dephosphorylated on treatment with the drug\textsuperscript{161}. Other studies have shown that the AKT is dephosphorylated on gefitinib treatment by the PI3K/AKT pathway. It has been suggested that the mutant EGFR use ERBB3 for the activation of PI3K. ERBB3 shRNA expressed in H3255 cells suppress the phosphorylation of AKT indicative of the role of ERBB3 in the PI3K-Akt pathway\textsuperscript{162}. Further, the SRC Family Kinases (SFKs) have been shown in another study to be constitutively activated in H3255 cell line further evidenced by SFK inhibitors inducing apoptosis in the cell line. Treatment with either PP1 or SKI-606 induced cleavage of PARP and caspase-3 and the cell line shifted to apoptosis mode indicative of the need of SFK phosphorylation for survival. NF\(\kappa\)B has also been shown to be hyperphosphorylated in H3255 from another study\textsuperscript{163} and the mechanism has been hypothesized to be that with activation of the EGFR receptor the I\(\kappa\)B\(\alpha\). Kinase kinase (IKKK) phosphorylates I\(\kappa\)B\(\alpha\) kinase signaling its degradation causing p50-p65 translocation activating NF\(\kappa\)B. The cell
line has also shown to express high levels of phosphorylated STAT3 and COX-2 in addition to EGFR but has low levels of E cadherin. A gefitinib or Erlotinib combination with Colexib reduced expression levels of COX-2, EGFR, p-EGFR, AKT and p-AKT levels. The caspases have known to play a very important role in triggering apoptosis in H3255. Levels of caspase-8 and caspase-9 are elevated nearly 2.5 fold 4 hours post Erlotinib treatment which in turn activates caspase-3. Time dependent exposure to Erlotinib induces caspase-8 to cleave BID to form truncated BID that translocates to the mitochondria to enhance mitochondrial mediated apoptotic pathways. Cytosolic levels of BAX increased post treatment and decreases later as the molecule translocated to the mitochondria. Erlotinib is known to induce conformational changes in BAX and BAK (whose levels remain the same post treatment in the cytosol and the mitochondria). This activation in independent of the ROS mediated pathways. Both Erlotinib and Gefitinib have similar inhibitory profiles and show same levels of apoptotic induction in the H3255 cell line and understanding the levels of proteins being effected on gefitinib treatment could also add on to information to Erlotinib based TKI data as well. H3255 was observed to more sensitive than other cells that were investigated with an IC50 of 40nM in a 72 hour survival assay. Treatment with 100mM gefitinib completely inhibited phosphorylation of EGFR in H3255. Downstream targets such as extracellular signal regulated kinase 1/2 (ERK 1/2) and also the v-akt murine thymoma viral oncogene homolog (AKT) were inhibited and other cell lines compared however showed phospho inhibition only when the concentration of the drug was increases a 100 times higher than what was used for H3255. The reason for such a behavior was attributed to the L858R mutation that was seen in H3255.