PROTEIN STRUCTURE DESIGNED USING AUTODOCK FOR A PHARMACOGENOMICS BASED DRUG FOR ASTHMA

D. Rachel Evangelene Tulip[1], K. Krishnadevi[2], K.K. Surya[2], K.Sekar[3]

[1]Department of Civil Engineerimg, Bharath University.Chennai-73

[2]Anna University, Guindy, Chennai-25

[3]Indian Institute of Science, Bangalore.

Abstract

. The β_2-adrenergic receptor is the most common adrenergic receptor in the lung, and associations between ADRB2 polymorphisms and intermediate phenotypes of asthma have been reported. The total number of SNPs in the ADRB2 gene was analyzed. According to the SNP profile. Using Modeller 9V2 software ADRB2 was modeled. Docking studies were performed with the available candidate drugs with the help of Autodock. The application of pharmacogenomics approach to Asthma will be essential for understanding the preventive mechanisms and could lead to individualized drug therapies in future.

1.Introduction:

In some individuals asthma is characterized by chronic respiratory impairment. In others it is an intermittent illness marked by episodic symptoms that may result from a number of triggering events, including upper respiratory infection, stress, airborne allergens, air pollutants (such as smoke or traffic fumes), or exercise. Some or all of the following symptoms may be present in those with asthma: dyspnea, wheezing, stridor, coughing, tightness and itching of the chest or an inability for physical exertion. Some asthmatics who have severe shortness of breath and tightening of the lungs never wheeze or have stridor and their symptoms may be confused with a COPD-type disease.

An acute exacerbation of asthma is commonly referred to as an asthma attack. The clinical hallmarks of an attack are shortness of breath (dyspnea) and either wheezing or stridor. Although the former is "often regarded as the sine qua non of asthma", some patients present primarily with coughing, and in the late stages of an attack, air motion may be so impaired that

no wheezing may be heard. When present the cough may sometimes produce clear sputum. The onset may be sudden, with a sense of constriction in the chest, breathing becomes difficult, and wheezing occurs (primarily upon expiration, but can be in both respiratory phases).

2.Materials and methods:

2.1MutDB

MutDB(http://mutdb.org/), to aid in determining which SNPs are likely to alter the function of their associated protein product. MutDB currently contains protein structure annotations and comparative genomic annotations for 8000 disease-associated mutations and SNPs found in the UC Santa Cruz (UCSC) Annotated Genome and the human RefSeq gene set. Normally multiple sequence alignments are used together in combination with protein structure to highlight functional mutations and functionally important protein regions.These data sets usually fall into two classes:Hand-curated, loci-specific databases that contain phenotypic annotations of rare variants and SNPs and Large unannotated data sets associated with many genes.

2.2 dbSNP

dbSNP Schema is very complex with well over 100 tables and many relationships among tables. One single ER (Entity Relationship) diagram with all dbSNP tables will be too huge to present useful information. Instead, separate tables can be made according to subject areas:Batch submission,Submitted SNP,population frequency and individual genotype. Frequency calculation by submitted SNP and population. SNP Mapping and Annotation.

2.3 Drugbank

The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. The database contains nearly 4800 drug entries including FDA-approved small molecule drugs, 123 FDA-approved biotech (protein/peptide) drugs, 71 nutraceuticals and >3,243 experimental drugs. Additionally, more than 2,500 non-redundant protein (i.e. drug target) sequences are linked to these FDA approved drug entries. Each DrugCard entry contains more than 100 data fields

with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data.

2.4 PharmGKB

PharmGKB is a publicly available Internet research tool developed by Stanford University with funding from the National Institutes of Health (NIH) and is part of the NIH Pharmacogenetics Research Network (PGRN), a nationwide collaborative research consortium. Its aim is to aid researchers in understanding how genetic variation among individuals contributes to differences in reactions to drugs.The PharmGKB database is a central repository for genetic, genomic, molecular and cellular phenotype data and clinical information about people who have participated in pharmacogenomics research studies. The data includes, but is not limited to, clinical and basic pharmacokinetic and pharmacogenomic research in the cardiovascular, pulmonary, cancer, pathways, metabolic and transporter domains. The contributers tab contains the links to all of the projects submitting data to the PharmGKB.

2.5 MODELLER

MODELLER is used for homology or comparative modeling of protein three-dimensional structures. The user provides an alignment of a sequence to be modeled with known related structures and MODELLER automatically calculates a model containing all non-hydrogen atoms. MODELLER implements comparative protein structure modeling by satisfaction of spatial restraints, and can perform many additional tasks, including de novo modeling of loops in protein structures, optimization of various models of protein structure with respect to a flexibly defined objective function, multiple alignment of protein sequences and/or structures, clustering, searching of sequence databases, comparison of protein structures, etc. MODELLER is available for download for most Unix/Linux systems, Windows, and Mac.

2.6Autodock

Autodock is a suite of automated docking tools. It is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D

structure.Autodock actually consists of two main programs: Autodock performs the docking of the ligand to a set of grids describing the target protein; AutoGrid pre-calculates these grids.In addition to using them for docking, the atomic affinity grids can be visualised. This can help, for example, to guide organic synthetic chemists design better binders.We have also developed a graphical user interface called AutoDockTools, or ADT for short, which amongst other things helps to set up which bonds will treated as rotatable in the ligand and to analyze dockings.

AutoDock has applications in:

- X-ray crystallography
- structure-based drug design
- lead optimization
- virtual screening
- combinatorial library design
- protein-protein docking
- Chemical mechanism studies

2.7 SIFT

SIFT is a sequence homology-based tool that Sorts Intolerant from Tolerant amino acid substitutions and predicts whether an amino acid substitution in a protein will have a phenotypic effect. SIFT is based on the premise that protein evolution is correlated with protein function. Positions important for function should be conserved in an alignment of the protein family, whereas unimportant positions should appear diverse in an alignment.

SIFT takes a query sequence and uses multiple alignment information to predict tolerated and deleterious substitutions for every position of the query sequence. SIFT is a multistep procedure that searches for similar sequences, chooses closely related sequences that may share similar function to the query sequence, obtains the alignment of these chosen sequences, and calculates normalized probabilities for all possible substitutions from the alignment. Positions with normalized probabilities less than 0.05 are predicted to be deleterious; those greater than or equal to 0.05 are predicted to be tolerated.

2.8 Swiss-model

SWISS-MODEL is a fully automated protein structure homology-modeling server, accessible via the ExPASy web server, or from the program DeepView (Swiss Pdb-Viewer). The purpose of this server is to make Protein Modelling accessible to all biochemists and molecular biologists World Wide.

2.9 SNPper

SNPper is a web-based application, developed in Common Lisp using the LispWeb development platform. It relies on a set of local databases containing information about genes and snps, and on remote access to the Draft Human Genome3 site to download up-to-date genomic sequences. The database currently contains information about 9,550 genes and over 1.2 million snps. Although the number of known snps is higher (currently close to 2 million) snpper only uses those whose exact position on a chromosome is known. snpper can be accessed at http://bio.chip.org/biotools/; its use is free for non-commercial research purposes.

2.10 BLAST

The Basic Local Alignment Search Tool (BLAST) finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families.

2.11 Genbank

GenBank is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences (Nucleic Acids Research, 2008 Jan;36(Database issue):D25-30). There are approximately 85,759,586,764 bases in 82,853,685 sequence records in the traditional GenBank divisions and 108,635,736,141 bases in 27,439,206 sequence records in the WGS division as of February 2008.

The complete release notes for the current version of GenBank are available on the NCBI ftp site. A new release is made every two months. GenBank is part of the International Nucleotide

Sequence Database Collaboration, which comprises the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.

2.12 PDB (Protein Data Bank)

The PDB archive contains information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies. As a member of the wwPDB, the RCSB PDB curates and annotates PDB data according to agreed upon standards.

The RCSB PDB also provides a variety of tools and resources. Users can perform simple and advanced searches based on annotations relating to sequence, structure and function. These molecules are visualized, downloaded, and analyzed by users who range from students to specialized scientists.

3.Result and Discussion:

3.1. Mutdb

Table 1: predicting the result of mutdb

Source ID	AA Position	WT->MT	Sequence	SIFT Score
DBSNP:rs1042713	15	G R	NCBI	Not scored
DBSNP:rs1042714	26	E Q	NCBI	Not scored
SWISS:VAR_003454	34	V M	Swiss-Prot	0.07
SWISS:VAR_009125	159	I F	Swiss-Prot	0.06
SWISS:VAR_009124	159	I L	Swiss-Prot	0.3
DBSNP:rs1800888	163	T I	NCBI	1.00
DBSNP:rs1042718	175	R G	NCBI	Not

				scored
DBSNP:rs3729943	219	S C	NCBI	0.24
SWISS:VAR_025101	220	S C	Swiss-Prot	0.16
SWISS:VAR_009394	375	K R	Swiss-Prot	0.25

3.2 dbSNP

Table 2: predicting the result of dbSNP

DbSNP cluster id	Function	DbSNP allele	Protein residue	Codon pos
	start codon			1
rs33973603	missense	G	Ser [S]	2
	contig reference	A	Asn [N]	2
rs1042713	missense	A	Arg [R]	1
	contig reference	G	Gly [G]	1
rs33957121	synonymous	T	His [H]	3
	contig reference	C	His [H]	3
rs35892629	synonymous	A	Gln [Q]	3
	contig reference	G	Gln [Q]	3
rs1042714	missense	C	Gln [Q]	1
	contig reference	G	Glu [E]	1

rs35336948	frame shift	C	Pro [P]	3
	contig reference	-		3
rs1042717	synonymous	A	Leu [L]	3
	contig reference	G	Leu [L]	3
rs35680672	frame shift	A	Tyr [Y]	2
	contig reference	-		2
rs1800888	missense	T	Ile [I]	2
	contig reference	C	Thr [T]	2
rs1042718	synonymous	A	Arg [R]	1
	contig reference	C	Arg [R]	1
rs3729943	missense	G	Cys [C]	2
	contig reference	C	Ser [S]	2
rs41320345	missense	C	Leu [L]	1
	contig reference	T	Phe [F]	1
rs41358746	missense	T	His [H]	3
	contig reference	G	Gln [Q]	3
rs56100672	missense	A	Arg [R]	1
	contig reference	G	Gly [G]	1
rs35933628	synonymous	T	Gly [G]	3
	contig reference	C	Gly [G]	3

rs1042719	synonymous	C	Gly [G]	3
	contig reference	G	Gly [G]	3
rs41354346	synonymous	C	Tyr [Y]	3
	contig reference	T	Tyr [Y]	3
rs3729944	synonymous	C	His [H]	3
	contig reference	T	His [H]	3
rs3730182	synonymous	C	Thr [T]	3
	contig reference	T	Thr [T]	3
rs1042720	synonymous	A	Leu [L]	3
	synonymous	T	Leu [L]	3
	contig reference	G	Leu [L]	3

The result obtained from the SNP tools, showed the presence of SNP in the gene ADRB2 at articular position also their prediction whether tolerated or deleterious .The information obtained from these tools helped to modeled the ADRB2 3D structure with SNP using modeller 9v6.

Position: 240

Change: F/L

MGQPGNGSAFLLAPNRSHAPDHDVTQQRDEVWVVGMGIVMSLIVLAIVFGNVLVITAI
AKFERLQTVTNY
FITSLACADLVMGLAVVPFGAAHILMKMWTFGNFWCEFWTSIDVLCVTASIETLCVIAV
DRYFAITSPFK

YQSLLTKNKARVIILMVWIVSGLTSFLPIQMHWYRATHQEAINCYANETCCDFFTNQAY
AIASSIVSFYV
PLVIMVFVYSRVFQEAKRQLQKIDKSEGR**L**HVQNLSQVEQDGRTGHGLRRSSKFCLKE
HKALKTLGIIMG
TFTLCWLPFFIVNIVHVIQDNLIRKEVYILLNWIGYVNSGFNPLIYCRSPDFRIAFQELLCL
RRSSLKAY
GNGYSSNGNTGEQSGYHVEQEKENKLLCEDLPGTEDFVGHQGTVPSDNIDSQGRNCST
NDSLL

Position: 247

Change: Q/H

MGQPGNGSAFLLAPNRSHAPDHDVTQQRDEVWVVGMGIVMSLIVLAIVFGNVLVITAI
AKFERLQTVTNY
FITSLACADLVMGLAVVPFGAAHILMKMWTFGNFWCEFWTSIDVLCVTASIETLCVIAV
DRYFAITSPFK
YQSLLTKNKARVIILMVWIVSGLTSFLPIQMHWYRATHQEAINCYANETCCDFFTNQAY
AIASSIVSFYV
PLVIMVFVYSRVFQEAKRQLQKIDKSEGRFHVQNLS**H**VEQDGRTGHGLRRSSKFCLKE
HKALKTLGIIMG
TFTLCWLPFFIVNIVHVIQDNLIRKEVYILLNWIGYVNSGFNPLIYCRSPDFRIAFQELLCL
RRSSLKAY
GNGYSSNGNTGEQSGYHVEQEKENKLLCEDLPGTEDFVGHQGTVPSDNIDSQGRNCST
NDSLL

Position: 257

Change: G/R

MGQPGNGSAFLLAPNRSHAPDHDVTQQRDEVWVVGMGIVMSLIVLAIVFGNVLVITAI
AKFERLQTVTNY

FITSLACADLVMGLAVVPFGAAHILMKMWTFGNFWCEFWTSIDVLCVTASIETLCVIAV
DRYFAITSPFK
YQSLLTKNKARVIILMVWIVSGLTSFLPIQMHWYRATHQEAINCYANETCCDFFTNQAY
AIASSIVSFYV
PLVIMVFVYSRVFQEAKRQLQKIDKSEGRFHVQNLSQVEQDGRTGH**R**LRRSSKFCLKE
HKALKTLGIIMG
TFTLCWLPFFIVNIVHVIQDNLIRKEVYILLNWIGYVNSGFNPLIYCRSPDFRIAFQELLCL
RRSSLKAY
GNGYSSNGNTGEQSGYHVEQEKENKLLCEDLPGTEDFVGHQGTVPSDNIDSQGRNCST
NDSLL

4 Structures of Normal and SNP Changed Proteins

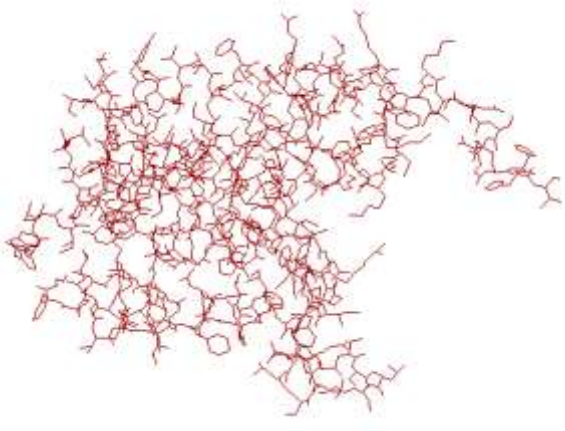

Figure 1: Structure of normal protein

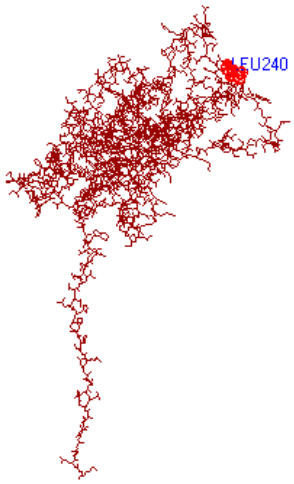

Figure 2:SNP in 240th position

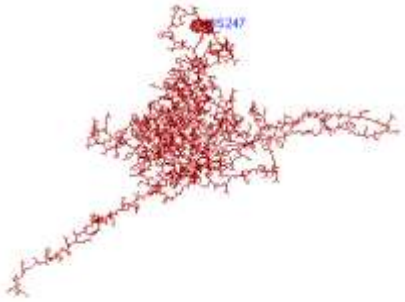

Figure 3: Snp in 247th position

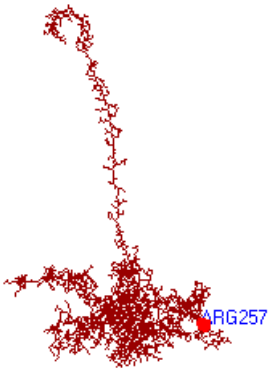

Figure 4: Snp in 257th position

3.3 Docking Results

Table 3: Estimated Free Energy of Binding for normal and SNP altered proteins

	Normal	**240**	**247**	**257**
Salbutemol	-2.94 kcal/mol	-5.86 kcal/mol	-4.75 kcal/mol	-3.16 kcal/mol

This table shows the comparative analysis of docking of two drugs with the normal protein and with the snp changed proteins.Here the normal represents the ADRB2 protein and the 240,247&257 represents the different SNP changed proteins.The number represents the position where the SNP change has been made.

4. CONCLUSION

SNPs do not have the nature to cause any disease, but they help in determining the likelihood that someone might develop a particular disease.

Docking studies were performed for both normal ADRB2 and SNP-ADRB2 with the available candidate drugs using Autodock. A drug (Salbutemol) and was taken and was docked with the normal protein and with the snp changed proteins. It has been found that

- Salbutemol can be suggested to patients having SNPs in 240,247,257 regions.

SNP profile based drug will reduce the toxicity and improve the capability of the treatement. The application of pharmacogenomics approach to Asthma will be essential for understanding the preventive mechanisms and could lead to individualized drug therapies in future.

5. References

1) Bach, P. B., Brown, C., Gelfand, S. E. and McCrory, D. C. (2001), "Management of acute exacerbations of chronic obstructive pulmonary disease: a summary andappraisal of published evidence". *Ann Intern Me.*, **134**, 600-620.

2) Ballester, E., Reyes, A., Roca, J., Guitart, R., Wagner, P, D. and Rodriguez-Roisin, R. (1989), "Ventilation-perfusion mismatching in acute severe asthma:effects of salbutamol and 100% oxygen", *Thorax.*, **44**, 258-267.

3) Barbera, J, A., Roca, J., Ferrer, A., Felez, M, A., Diaz, O, Roger, N. and Rodriguez-Roisin, R. (1997), "Mechanisms of worsening gas exchange during acute exacerbations of chronic obstructive pulmonary disease", *Eur Respir J.*, **10**, 1285-1291.

4) Bateman ED, Bousquet J, Braunstein GL. Is overall asthma control being achieved? A hypothesis-generating study. Eur Respir J 2001;17:589-595.

5) Bateman ED, Boushey HA, Bousquet J, Busse WW, Clark TJ, Pauwels RA, Pedersen SE, and the GOAL Investigators Group. Can guideline-defined asthma control be achieved? The Gaining Optimal Asthma ControL study. Am J Respir Crit Care Med 2004;170:836-844.

6) Bernasconi, M., Brandolese, R., Poggi, R., Manzin, E. and Rossi, A.(1990), "Dose-response effects and time course of effects of inhaled fenoterol on respiratory mechanics and arterial oxygen tension in mechanically ventilated patients with chronic airflow obstruction", *Intensive Care Med.*, **16**, 108-114.

7) Cardus, J., Burgos, F., Diaz, O., Roca, J., Barbera, J, A., Marrades, RM., Rodriguez-Roisin, R. and Wagner PD. (1997), "Increase in pulmonary ventilation-perfusion inequality with age in healthy individuals", *Am J Respir Crit Care Me.,* **156**, 648-653.

8) Celli, B, R. and MacNee, W. (2004), "Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper", *Eur Respir J*, **23.**, 932-946.

9) Diaz, O., Barbera, J, A., Marrades, R., Chung, K, F., Roca, J. and Rodriguez-Roisin, R. (1997), "Inhibition of PAF-induced gas exchange defects by beta-adrenergic agonists in mild asthma is not due to bronchodilation", *Am J Respir Crit Care Me.,* **156**, 17-22.

10) Gabrijelcic, J., Casas, A., Rabinovich, R, A., Roca, J., Barbera, J,A., Chung, K, F. and Rodriguez-Roisin, R. (2004), "Formoterol protects against platelet-activating factor-induced effects in asthma", *Eur Respir J.,* **23**, 71-75.

11) Gale, G, E., Torre-Bueno, J, R., Moon, R, E., Saltzman, H, A. and Wagner, P, D. (1985), "Ventilation-perfusion inequality in normal humans during exercise at sea level and simulated altitude", *J Appl Physio.,* **58**, 978-988.

12) Karpel, J, P., Pesin, J., Greenberg, D. and Gentry, E. (1990), "A comparison of the effects of ipratropium bromide and metaproterenol sulfate in acute exacerbations of COPD", *Chest.*, **98**, 835-839.

13) Kessler, R., Faller, M., Fourgaut, G., Mennecier, B. and Weitzenblum, E. (1999), "Predictive factors of hospitalization for acute exacerbation in a series of 64 patients with chronic obstructive pulmonary disease", *Am J Respir Care Med*, **159**, 158-164.

14) McDonald, D, M. (1990), "The ultrastructure and permeability of tracheobronchial blood vessels in health and disease", *Eur Respir J Suppl.*,**12**, 572-585.

15) Mendes, E, S., Campos, M, A. and Wanner, A. (2006), "Airway blood flow reactivity in healthy smokers and in ex-smokers with or without COPD", *Chest.*, **129**, 893-898.

16) Parker, C, M., Voduc, N., Aaron, S, D., Webb, K, A. and O'Donnell, D, E. (2005), "Physiological changes during symptom recovery from moderate exacerbation of COPD", *Eur Respir J.*, **26**, 420-428.

17) Polverino, E., Gómez, F, P., Rodríguez, A, C., Peña, A., Valera, J, L., Soler, N., Barberà, J, A., Roca, J, R. and Rodríguez-Roisin, R. (2005), "Nebulized salbutamol transiently deteriorates pulmonary gas exchange in COPD exacerbations", *Proc Am Thorac Soc*, **2**, 406-408.

18) Polverino, E., Gómez, F, P., Valera, J, L., Soler, N., Roca, J., Barberà, J, A. and Rodriguez-Roisin, R. (2006), "Gas exchange response to nebulized salbutamol in exacerbations and stable COPD", *Proc Am Thorac So,*. **3**, 721-725.

19) Rabe, K, F., Hurd, S., Anzueto, A., Barnes, P, J., Buist, S, A., Calverley, P., Fukuchi, Y., Jenkins, C., Rodriguez-Roisin, R. and van Weel, C. (2006), "Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease", *Eur Respir J.*, **26**, 480-488.

20) Ringsted, C, V., Eliasen, K., Andersen, J, B., Heslet, L. and Qvist, J. (1989), "Ventilation-perfusion distributions and central hemodynamics in chronic obstructive pulmonary disease:effects of terbutaline administration", *Chest.*, **96**, 976-983.

21) Roca, J., Burgos, F., Barbera, JA., Sunyer, J., Rodriguez-Roisin, R., Castellsague, J., Sanchis, J., Anto, J, M., Casan, P. and Clausen, J, L. (1998), "Prediction equations for plethysmographic volumes", *Respir Me*, .**92**, 454-460.

22) Roca, J. and Wagner, P, D. (1994), "Contribution of multiple inert gas elimination technique to pulmonary medicine. 1: principles and information content of the multiple inert gas elimination technique", *Thora.*, **49**, 815-824.

23) Roca, J., Burgos, F., Sunyer, J., Saez, M., Chinn, S., Anto, J, M., Rodriguez-Roisin, R., Quanjer, P, H., Nowak, D. and Burney, P. (1998), "References values for forced spirometry: Group of the European Community Respiratory Health Survey", *Eur Respir.*, **11**, 1354-1362.

Rodriguez-Roisin, R. (2000), "Toward a consensus definition for COPD exacerbations", *Chest.*, **117**, 398S-401S.

Rodriguez-Roisin, R., Ballester, E., Roca, J., Torres, A. and Wagner, P, D. (1989), "Mechanisms of hypoxemia in patients with status asthmaticus requiring mechanical ventilation", *Am Rev Respir Di.,* **139**, 732-739.

Rodriguez-Roisin, R. (2006), "COPD exacerbations.5: management", *Thorax,* **61**, 535-544.

Stevenson, N, J., Walker, P, P., Costello, R, W. and Calverley, P, M. (2005), "Lung mechanics and dyspnea during exacerbations of chronic obstructive pulmonary disease", *Am J Respir Crit Care Med.*, **172**, 1510-1516.

Wagner, P, D. and Rodriguez-Roisin, R. (1991), "*Clinical advances in pulmonary gas exchange*", *Am Rev Respir Dis.,* **143**, 883-888.