

Mapping antigenic diversity and strain specificity of mumps virus: A bioinformatics approach

Urmila Kulkarni-Kale ^{a,*}, Janaki Ojha ^a, G. Sunitha Manjari ^a, Deepti D. Deobagkar ^b,
Asha D. Mallya ^c, Rajeev M. Dhere ^c, Subhash V. Kapre ^c

^a *Bioinformatics Centre, University of Pune, Pune 411007, India*

^b *Department of Zoology, University of Pune, Pune 411007, India*

^c *Serum Institute of India Ltd., 212/2, Hadapsar, Pune 411028, India*

Received 13 July 2006; returned to author for revision 18 August 2006; accepted 15 September 2006

Available online 1 November 2006

Abstract

Mumps is an acute infectious disease caused by mumps virus, a member of the family *Paramyxoviridae*. With the implementation of vaccination programs, mumps infection is under control. However, due to resurgence of mumps epidemics, there is a renewed interest in understanding the antigenic diversity of mumps virus. Hemagglutinin–neuraminidase (HN) is the major surface antigen and is known to elicit neutralizing antibodies. Mutational analysis of HN of wild-type and vaccine strains revealed that the hypervariable positions are distributed over the entire length with no detectable pattern. In the absence of experimentally derived 3D structure data, the structure of HN protein of mumps virus was predicted using homology modeling. Mutations mapped on the predicted structures were found to cluster on one of the surfaces. A predicted conformational epitope encompasses experimentally characterized epitopes suggesting that it is a major site for neutralization. These analyses provide rationale for strain specificity, antigenic diversity and varying efficacy of mumps vaccines.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Mumps virus; Hemagglutinin–neuraminidase; HN; Wild and vaccine strain; Homology model; Sequential epitope; Conformational epitope; Bioinformatics; Molecular immunology

Introduction

Mumps is an acute infectious viral disease characterized by enlargement of the parotid and salivary glands. Other complications of mumps include permanent deafness, orchitis, pancreatitis and aseptic meningitis (AM). The virus usually spreads through respiratory droplets and humans are the only natural host although non-human primates, rodents and other species can be experimentally infected (Wolinsky, 1996).

Mumps virus is a member of the family *Paramyxoviridae*, genus *Rubulavirus* (Rima et al., 1995). The virus is enveloped and its 15.3-kb genome is non-segmented, single-stranded, negative-sense RNA. It codes for seven proteins viz., nucleocapsid (N), phosphoprotein (P), matrix (M), fusion (F),

small hydrophobic (SH), hemagglutinin–neuraminidase (HN) and large (L) proteins. Besides these, the phosphoprotein gene (P) also codes for two more proteins, V and I (Paterson and Lamb, 1990). HN is the major antigenic protein, known to elicit neutralizing antibodies. The SH gene is known to be the most variable part of the genome and phylogenetic reconstruction studies of different isolates of mumps virus using SH genes revealed the existence of 12 genotypes, designated A to L and 3 unnamed potential new genotypes (Jin et al., 2005; Tecle et al., 2001).

Different genotypes have been shown to co-circulate (Afzal et al., 1997a, 1997b; Wu et al., 1998; Tecle et al., 2001, 2002; Takahashi et al., 2000) and their distribution may vary among closely related regions within a country (Takahashi et al., 2000; Tecle et al., 2001). Strains of mumps virus are also known to exhibit varying degrees of neurovirulence (Merz and Wolinsky, 1981; Saito et al., 1996; Rubin et al., 1998, 2000; Rafiefard et al., 2005; Sauder et al., 2006). However, the relative degree of

* Corresponding author. Fax: +91 20 2569 0087.

E-mail addresses: urmila@bioinfo.ernet.in,
urmila.kulkarni.kale@gmail.com (U. Kulkarni-Kale).

their neurovirulence could not be determined mainly due to lack of animal models. Preliminary results obtained using neonatal rat (Rubin et al., 2000, 2005) and Marmoset monkey models (Saika et al., 2004) for vaccine strains, though promising, require further investigations.

Live attenuated mumps vaccines are available as monovalent mumps, bivalent measles–mumps (MM) and trivalent measles–mumps–rubella (MMR) vaccines. The strains viz., Jeryl Lynn (JL), Leningrad-3 (L-3), L-Zagreb, Rubini and Urabe have been in use all over the world since early 1980s and have been reviewed extensively (Galazka et al., 1999; Furesz, 2002; Folb et al., 2004; Ivancic et al., 2005). It has been reported that the JL vaccine is a mixture of two strains, JL-2 (also called minor) and JL-5 (also called major) (Afzal et al., 1993; Amexis et al., 2002). Similarly, the Urabe vaccine strain has also been reported as a mixture of strains (Brown et al., 1996). The Rubini strain does not appear to provide long-term protection (Galazka et al., 1999; Utz et al., 2004; Ong et al., 2005). While offering varying degrees of protection (Ong et al., 2005), most of the vaccine strains are known to cause adverse reactions such as AM (Flynn and Mahon, 2003; Folb et al., 2004; Nagai et al., 2006). Another challenge is co-circulation of wild-type virus SBL-1 while mumps vaccine is in use (Orvell et al., 1997; Tecle et al., 1998).

Mumps infection is under control due to the implementation of vaccination programs. However, there is a renewed interest to understand the antigenic diversity of mumps virus because of recent outbreaks of mumps epidemics (CDC, 2006). The knowledge thus gained will play a decisive role in mumps vaccine development (Amexis et al., 2001; Furesz, 2002). The contribution of specific humoral response to mumps virus as a defense factor has not been definitively explained (Pipkin et al., 1999; Kacprzak-Bergman et al., 2001). However, HN protein is a major target for humoral immune response in mumps virus infection as it elicits neutralizing antibodies (Tanaka et al., 1992; Cusi et al., 2001). The regions viz., 265–288, 329–340 and 352–360 of HN have been reported to evoke immune response and are also responsible for virulence (Kovamees et al., 1990; Orvell et al., 1997; Cusi et al., 2001). Furthermore, the region 329–340 is shown to have the ability to induce neutralizing antibodies not only to the attenuated virus strains but also to wild-type strains (Cusi et al., 2001).

Nucleotide and protein sequences of HN from a number of mumps virus strains/isolates are available in the public domain repositories. Extensive analyses of these sequences helped in identification of strain-specific variations and characterization of mutants (Yates et al., 1996). Attempts have also been made to correlate mutations with properties such as antigenicity and neurovirulence (Sauder et al., 2006). It has been reported that mutations in HN are distributed over the entire length and no pattern could be detected using sequence data alone (Ivancic et al., 2005).

It is known that the patterns, which are apparently hidden at sequence level, become evident when mapped onto structure. Experimentally derived 3D structure data of HN protein of mumps virus are not available. Therefore, an attempt has been made to predict its 3D structure using knowledge-based homology modeling approach. The structures of HN protein

of one wild-type (SBL-1) and three vaccine strains, namely JL2, JL5 and L-Zagreb were predicted. The observed variations of amino acids among the groups of vaccine and wild-type strains were then mapped on the predicted structure. Sequential and conformational epitopes were predicted and analyzed in the context of observed mutations (Kulkarni-Kale et al., 2005; Kolaskar and Kulkarni-Kale, 1999; Kolaskar and Tongaonkar, 1990).

Results

Compilation of HN sequences

A total of sixty-four nucleotide sequences of HN from various strains/isolates of mumps virus were retrieved from GenBank. It must be mentioned that curation of data is an important step in any Bioinformatics analyses. Although information is available, it is scattered either in different entries of the same database or among various databases. For example, only a few sequence entries for HN in GenBank are annotated with respect to genotype data. Since the genotyping of mumps virus is carried out using SH gene, the sequence entries of SH for a given strain/isolate were searched to retrieve the genotype information. However, only 49 out of 64 strains could be annotated with respect to genotype information using this approach.

Multiple sequence alignment

The nucleotide and protein sequences of HN were aligned using ClustalW (Chenna et al., 2003). As expected, though there exists sequence similarity among the strains of mumps virus, a few strain-specific variations are observed. Multiple sequence alignment of HN protein showed 90% identity (marked with * in the MSA) and 96% similarity (identity + favorable substitutions that are marked with : and . in the MSA) among vaccine strains (see Annexure V). These values varied up to 74% and 91% respectively in wild-type strains (MSA data not shown). The known motifs viz., leucine-zipper, neuraminidase (240-NRKSCS-245) and receptor-binding site of hemagglutinin (405-GAEGRV-410) are conserved among all 64 entries (Jorgensen et al., 1987; Mirza et al., 1994; Lim et al., 2003). Similarly, there are 9 potential N-linked glycosylation sites with signature sequence N-X-S or N-X-T (Apweiler et al., 1999) at positions: 12, 127, 284, 329, 400, 448, 464, 507 and 514. Of these, glycosylation sites at positions 127, 284, 448, 507, and 514 are conserved in both vaccine and wild-type strains. Mutation of N to D/S at position 12 results in loss of a potential cytosolic glycosylation site in a few strains. This site was found to be missing in the isolates of mumps virus from a vaccinated population in Singapore (Lim et al., 2003). The 3rd amino acid position in the glycosylation site 329 contains either, T or S. However, at similar position in glycosylation site 400, I is found in entries with GenBank accession numbers: AF448531, AF448530, AF448527, AF448534 and AF448528, an unfavorable mutation as far as glycosylation is concerned. Such strains may not get glycosylated and this may account for further

strain-specific antigenic variation, brought in due to post-translational modifications. Similarly, the N at position 464 is substituted by K/H, which would result in loss of this glycosylation site. The mutation N464K is observed in wild-type strains RW (genotype D: GenBank accession no. M19933) and 871005 (genotype B: GenBank accession no. AF314562), as well as in vaccine strains Smith-Kline Beecham live-attenuated strain (genotype B: GenBank accession no. AF314559) and JL2 (genotype A: GenBank accession no. AY584604, AF345290). The mutation N464H is observed in JL5 strains (GenBank accession no. AY584603 and X93179). In addition to mutations in the glycosylation site, a few escape mutants have been characterized by several workers worldwide. These include positions 239, 266, 269, 329, 352, 354, 356 and 360, which are identified using SBL-1 and Kilham strains (Orvell et al., 1997; Rafiefard et al., 2005). The MSA data and the literature-curated data on known escape mutants were subsequently used to map molecular determinants of antigenicity.

In the next phase, the sequences of all known wild-type and vaccine strains were aligned separately. These MSA revealed that there are 44 and 23 mutations respectively in wild-type and vaccine strains. Of these, 15 sites were found to be common between both the groups (Table 1). The variability index (VI) calculated using the method of Wu and Kabat (1970) for wild-type and vaccine strains are shown as Annexures I and II respectively. VI gives an idea about the extent of variation at a given position among a group of aligned sequences and is the ratio of the number of different amino acids at a given position and frequency of the most common amino acid at that position. Eight sites show VI >4 among wild-type strains as compared

with only three sites in vaccine strains (Annexures I and II). The eight variable sites of wild-type strains (12, 20, 25, 121, 266, 354, 385, 533), when mapped on the predicted structure, revealed that three sites (12, 20, 25) are part of cytosolic region; site 121 is not modeled; site 385 is buried and 3 sites (266, 354 and 533) are present on the surface of HN and are accessible for interactions with antibodies. The positions 266 and 354 are of significance since they are part of the experimentally characterized neutralizing epitope. Among the vaccine strains, only three sites: 8, 80 and 121 are found to have variability index of 4 or higher. Of these sites, 8 and 80 are part of cytosolic region; site 121 is found to be hypervariable in both vaccine and wild-type strains and is a part of the stalk region. It must be mentioned that 3D structures are predicted only for the extracellular domain of HN protein.

Prediction of 3D structure

Homology modeling is the method of choice for the prediction of 3D structure of proteins when structure of at least one orthologue is available and the sequence similarity between the query and its orthologue is $\geq 30\%$ (DeWeese-Scott and Moul, 2004). The 3D structures of HN of one wild-type (SBL-1) and three vaccine strains (viz. JL2, JI5 and L-Zagreb) were predicted using the said approach. The choice of SBL-1, a wild-type virus, was based on two reasons: firstly, SBL-1 is less neurovirulent than other wild-type strains; and secondly it has been reported to co-circulate among the populations vaccinated with JL vaccine (Orvell et al., 1997; Teclé et al., 1998). The 3D structure data of HN of two paramyxoviruses viz., parainfluenza virus (PIV) and Newcastle disease virus (NDV) are available in Protein Data Bank (PDB). The HN sequences of the four mumps virus strains mentioned above were aligned with their orthologues in PDB to identify template(s) for homology modeling. The sequence of HN of mumps virus was found to have 35% and 27% similarity with NDV and PIV respectively. On the basis of sequence similarity, stereochemical and geometrical evaluation of 3D structures of HN of NDV and PIV, the structure of HN of NDV was identified as a template (PDB ID: 1E8U; Crennell et al., 2000; Takimoto et al., 2000). The pairwise sequence alignment of HN of mumps virus (SBL-1) with its template is shown in Fig. 1. The initial structure of HN of SBL-1 was then built by assigning the coordinates from the aligned residues of template. Care was taken to avoid assignment of initial coordinates to Proline from non-Proline residues and to non-Glycine from Glycine residues. Coordinates for the loop regions were assigned by searching the database of loops (Hobohm and Sander, 1994). Energy of the model is calculated using amber all atom force field (Seibel et al., 1990) and distance-dependent dielectric constant ($4r_{ij}$). This helps to simulate the effect of the solvent in an implicit fashion. The bond lengths, bond angles and omega values were checked and corrected only if they were outside the acceptable range. For bond length and bond angle, the acceptable range was fixed as (standard bond length/bond angle $\pm 3\sigma$) and for $\omega = 180^\circ \pm 15^\circ$.

Table 1
Analysis of mutations in HN protein of mumps virus

Sites	Mutations: in vaccine strains, derived using MSA of vaccine strains	Mutations: in wild-type strains, derived using MSA of wild-type strains
Informative and variable	V135I , <u>V212I</u> , V218A , V259I, D266A , T279I , V287I , T288K , <u>S295T</u> , S336L , Q354P , E356D , S372N , S442Y , S462L , N464K/H , E468K , V470I , T473I , V474A , S490R , <u>T533N</u> , <u>T577A</u> ²³⁽¹⁵⁾	N131S, I132V, D156E, H161R , K203R/N, V212I, N240I, D251G, D266E/A/N , <u>T279I</u> , V287I , S295T/A, K335R, S336L , Y347C, S351L/P, P353S, Q354P/K/H , D356E/S , R363K, F370L, N372S , V375I, A378V, P397A, S398L, N399S, T402I, L403M, L411I, I435V/L, T438I, Q444L/P, V447G, N464K , E468K , T473I , A474V , P483S, H491Y, T511A/N, <u>T533N/A</u> , T577A , T582I/N 44(15)

Note. Mutations only in the extracellular domain, i.e., residues 132–582 are listed for both wild-type and vaccine strains. The mutations are reported in the format ‘predominant residue–position–mutated residue’. Sites common in vaccine and wild-type strains are underlined. Sites mapped on predicted structure (see Fig. 2b) are shown in bold.

```

1E8U_A: GAPIHDPDFIGGIGKELIVDNAS-DVTSFYPSAFQEHLENFIPAPITGSG
SBL-1: NIPLVNDRFINGINKFIIEDYATHDFISIGNPL---NMPSFIPATATSPNG

1E8U_A: CTRIPSFDMASATHYCYTHNVILSGCRDHSKSHQYALGLVLRITATGRIFF
SBL-1: CTRIPSFSLGKTHWCYTHNVINANCKDHTSSNQYVSMGILVQTASGYPMF

1E8U_A: STLRISLDDTQNRKSCSVSATPLGCDMLCSKVTETEEEDYNSAVPTLMA
SBL-1: KTLKIQYLSLGLNRKSCSIATVPDGCAMCYVSTQLEANDYAGSSPPTQK

1E8U_A: HGRLEGDFDQYHEKDLVDTLTFEDWVANYPGVGGGSFIDGRVWFVSVYGGK
SBL-1: LTLFLYNDTITERTISPSGLEGNWATLVPGVGGSGIYFENKLIFFPAYGGVL

1E8U_A: PNSPSDVTQEGKVIYIKRYNDTCPDEQDYQIRMAKSSYKPGRFGGKRIQQ
SBL-1: PNSTLVGKSAREFFRVNPNPCSGPPQELDQRALRSYFPRYFSSRRVQS

1E8U_A: AILSIKIVSTSLGEDPVLTVPPNTVTLMGAEGRILTVGTSHFYLYQRGSSYF
SBL-1: AFLVCAWNQILLVTNCELVVPNNQTLMGAEGRVLLINNRLLYYQRSTSWW

1E8U_A: SPALLYPMTVSN---KTATLHSPYTFNAFTRPGSIPCQASARCNSCVT
SBL-1: PYELLYEISFTFTNSGQSSVNMNWIPIYSFTTPGSGNCSGKNVCPTVCVS

1E8U_A: GVYTDYPLIFYR--NHTLRGVFGTMLD--SEQARLNPAVAEFDSTSRSR
SBL-1: GVYLDWPLTPYSHQSGINRNRYFTGALLNSSTTRVNPTLYVSALNNLKV

1E8U_A: ITRVSSSSTKAAAYTTSTCFKVVKTNKTYCLSLIAEISNTLFGFRIVPLLV
SBL-1: LAPYGTQGLFASYTTTTCFQDTGDASVYCYVIMELASNIVGEFQILPVLA

1E8U_A: EILKND
SBL-1: RLTTT

```

Fig. 1. Pairwise sequence alignment of HN (monomer) of mumps virus (query: SBL-1) and that of NDV (template: 1E8U; chain A). The structurally conserved regions (SCRs) and loop regions are identified by superimposition of the structures of HN proteins and boundaries of secondary structural elements of NDV. SCRs are shown as highlighted text.

Models of all four strains were built and optimized independently to eliminate bias. The structure of the monomer for every strain was built initially followed by dimer assembly and optimization. The order for optimization was: loops, SCRs, all side-chains and all atoms.

Evaluation of predicted structures

The essential accuracy and correctness of the models were evaluated using PROSTAT (module in Homology), ProsaII (Sippl, 1993) and PROCHECK (Laskowski et al., 1993). Evaluation of the models using more than one method was essential in this particular case because the sequence identity between the HN protein of mumps virus and NDV is only 35%. The models were evaluated in terms of stereochemical and geometrical parameters like bond lengths, bond angles, torsion angles and packing environment and were found to satisfy the said criteria (Ramachandran and Sasisekharan, 1968) (Table 2). It must be mentioned that the accuracy of any model built using homology modeling is limited not only by the accuracy of the template but also sequence alignment, choice of initial structures of loop regions and extent of optimization (Rost and O'Donoghue, 1997).

Further objective check on the quality of models was obtained using the program ProsaII (Sippl, 1993). The energy

graphs drawn using ProsaII display the energetic architecture of protein folds as a function of amino acid sequence. The ProsaII Z-score for the template and the four models are found to be in the range typical for globular proteins (Table 2). The PROCHECK G factor is the log-odds score based on the observed distributions of various stereochemical parameters (Laskowski et al., 1993). The over all G factor score for the template and models lie in the range -0.5 to 0.5 which indicates that all four structures are stereochemically acceptable (Table 2).

Structural description of HN of mumps virus

Predicted structure of HN of SBL-1 is shown in Fig. 2a. As can be seen, the predominant secondary structure is β -sheet. The fold of the protein is β -propeller. Every monomer is a six-bladed β -propeller consisting of six 4-stranded β -sheet motifs along with four helices. Superimposition of the predicted structure of HN of all strains on the template structure shows very high similarity at 3D structure level, though sequence similarity is only 35%. The root-mean-square (rms) deviations (\AA) between the template and each of the models for ~ 428 structurally aligned residues are SBL-1: 0.63, J12: 0.67; J15: 0.65 and L-Zagreb: 0.67, respectively. Furthermore, all the four helices and majority of the strands are conserved and show similar boundaries. However, some variations in the length of the strands have been observed when compared with the template structure. A few short strands that are present in the template have disappeared in the model. Similarly, a few strands are split into two short strands. The rmsd among the models of HN ranged between 0.37 and 0.5 \AA indicating that there exists an overall structural similarity for the backbone atoms. Subsequent to our homology modeling studies, the coordinates of SV-5, a member of *Rubulavirus* genus, were available in PDB (Yuan et al., 2005). Since SV-5 is evolutionarily closely related to mumps virus as compared to NDV (chosen template), we compared the crystal structures of HN of SV-5 (PDB: 1Z4V) and NDV (PDB: 1E8U). The overall structural rmsd between 1Z4V and 1E8U (for a monomer) was found to be 0.94 \AA

Table 2
Evaluation of predicted structures of HN monomer of mumps virus

	PROCHECK G factor (Laskowski et al., 1993)	ProsaII Z-score (Sippl, 1993)	Occupancy in Ramachandran plot (Ramachandran and Sasisekharan, 1968) (%)	Residues in disallowed regions
Template	0.22	-9.37	99.4	I127, F156, D198, T551
SBL-1	-0.16	-7.9	98.9	P133, V135, K202, D285, L506
Lzagreb	-0.18	-8.01	99	V135, L163, E315, S490
JL2	-0.23	-7.69	99	V135, E315, V389, L506
JL5	-0.18	-8.01	98.9	E315, N391, N415, R490

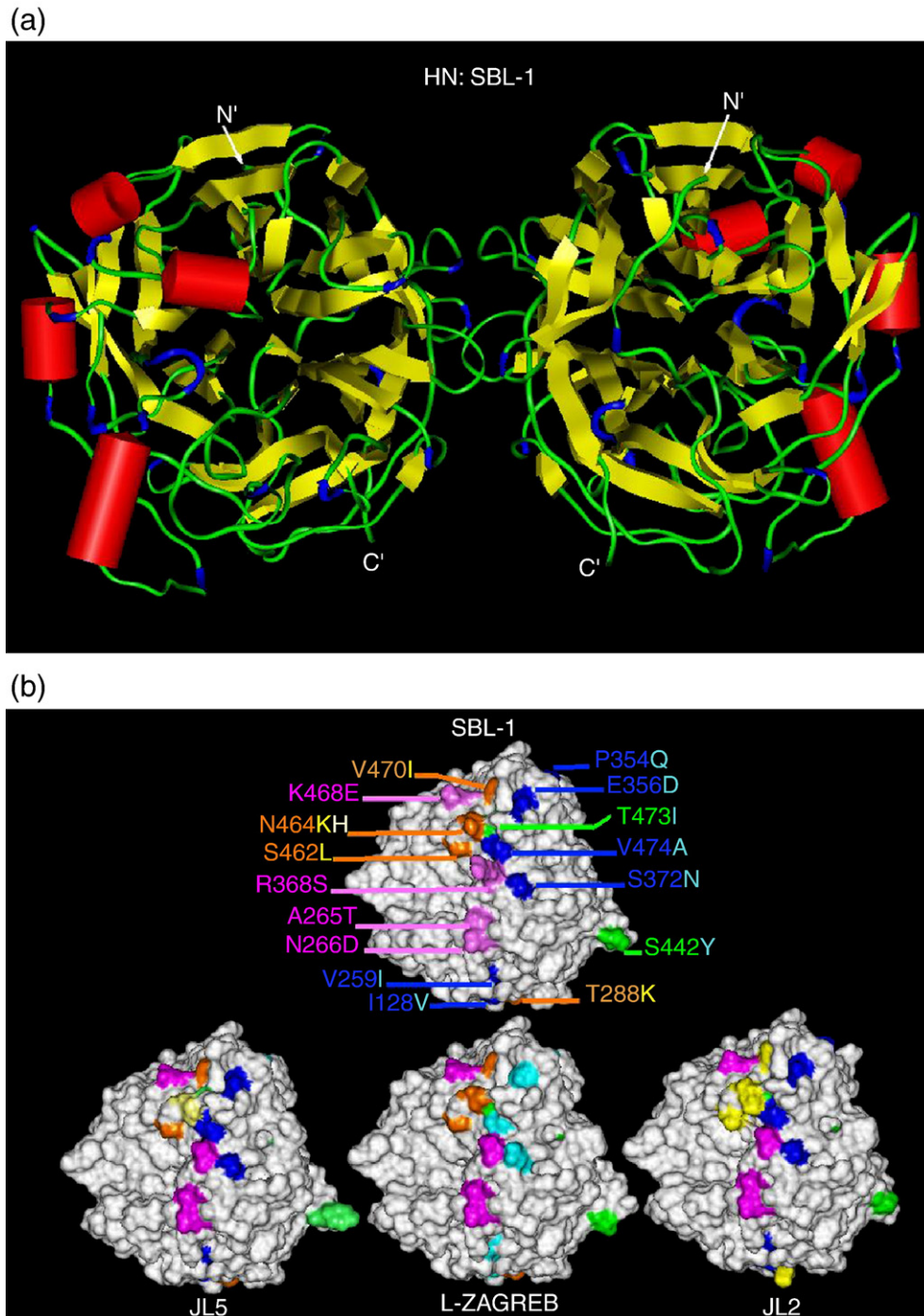


Fig. 2. (a) Predicted structure of HN (dimer) of L-Zagreb. Kabsch and Sander secondary rendering is shown. The β -sheets form the major secondary structure and four short α -helices are present in each chain. The strands, helices, turns and coils are shown in yellow, red, blue and green respectively. (b) Mutations mapped on the solvent accessible surface of the predicted structures of HN (monomer) of SBL-1, JL5, JL2 and L-Zagreb. Mutations are colored according to predominance of a given amino acid. Dark color indicates that the said amino acid is present in 3 of the four strains.

indicating a similar fold. The structural superimposition between four models of HN of mumps virus (one wild-type and three vaccine strains) built using 1E8U as a template and 1Z4V showed a deviation in the range of 0.85–0.88 Å. Thus, comparison with orthologue further validated the predicted structures.

The predicted structure of SBL-1 was compared with that of vaccine strains to study the effect of mutations. The comparison

of SBL-1 and L-Zagreb structures shows that the mutations P₄₅₉R, V₂₅₉I, A₅₇₇T that are part of SCR regions have effect on local region rather than on surrounding regions. However, in case of N₁₆₁H mutation, more effect is seen in surrounding regions rather than in local regions as it is a part of splice site connecting two SCRs. Some of the mutated residues are part of predicted antigenic determinants, which may have implications in antibody recognition.

Mapping of epitopes

The regions, 265–288, 329–340 and 352–360, of HN are known to be antigenic (Kovamees et al., 1990; Orvell et al., 1997; Cusi et al., 2001). Mapping of these regions on the predicted structure of mumps revealed that they are present on the surface of HN and hence are available for interaction with antibodies. The B cell epitopes were predicted using both, sequence data and 3D models of HN. Some of the residues predicted as antigenic using sequence-based approach are buried in the core (Kolaskar and Tongaonkar, 1990). This is due to the fact that sequence-based approaches use only the amino acid propensity data for calculation of solvent accessibility. The epitopes were also predicted using CEP server, a structure-based epitope prediction program (Kulkarni-Kale et al., 2005; Kolaskar and Kulkarni-Kale, 1999) that predicts both, conformational and sequential epitopes (see Supplementary material: Annexures III and IV) using accessibility of amino acids in an explicit manner. The conformational epitopes or probable antibody-binding sites are composed of multiple individual sequential epitopes that are brought together due to protein folding (Van Regenmortel, 1998). The accuracy of the Conformational Epitope Prediction algorithm was found to be 76% when evaluated using the co-crystal structures of antigen–antibody complexes deposited in PDB. The algorithm was also found to correctly predict one of the sequential epitopes of the hyaluronidase, a major allergen of bee venom that has been characterized using the Fab fragment of a mouse monoclonal IgG antibody. Out of 14 residues predicted by CEP server, 9 sequential residues are involved in Fab binding (Zora Housley, personal communication). The CEP server predicted 20 antigenic determinants (see Supplementary material: Annexures III and IV) that form 16 conformational epitopes and 3 sequential epitopes. The predicted B cell epitopes, 261–266, 269–272, 284–296, 327–331 and 334–363, overlap with the experimentally characterized epitopes of mumps virus. Furthermore, one of the predicted epitope (CE5) comprising of regions 261–266 and 269–272, has already been validated to be a neutralizing epitope (Kovamees et al., 1990; Orvell et al., 1997). Similarly, CE4 is made up of regions 261–266, 269–272, 199–207 and 220–240 suggesting that these regions form a major antibody-binding site, capable of being recognized by multiple antibodies. Antibodies with overlapping binding sites have been well characterized for lysozyme (Davies and Cohen, 1996; Smith-Gill, 1996). The region 213–372 has been reported to induce hemagglutination-inhibiting and mumps virus neutralizing antibodies (Cusi et al., 2001). As can be seen from the supplementary material (Annexures III and IV), CEP predicts six epitopes in the region 213–372. It has also been reported that the antibodies raised against the synthetic peptide 329-NSTLGVSAREF-340 neutralized the mumps wild-type virus. However, these antibodies failed to neutralize the Urabe Am-9 vaccine strain (Cusi et al., 2001). It must be mentioned that Urabe AM-9 was found to have K335E mutation, which would require charge reversal in the CDR regions of respective antibodies. The lower efficacy of Urabe AM-9 could be attributed to such mutations (Ong et al., 2005). Two epitopes

327–331 and 334–359 that encompass the peptide 329-NSTLGVSAREF-340 are predicted by the CEP server (Annexure IV). It was observed that the region 334–359 extends up to 363 when predicted for JL2, JL5 and SBL-1 (data not shown).

Discussions

Hemagglutinin–neuraminidase (HN) protein has been identified as a major target for the humoral immune response in mumps virus infection (Cusi et al., 2001). Nucleotide sequences of HN of 64 strains/isolates of mumps virus are available in GenBank release 146 (Feb. 2005). Most of these were sequenced to characterize the outbreaks in the vaccinated populations and to identify the spread and circulation of genotypes (Teclé et al., 2001; Lim et al., 2003).

The complete genomes of mumps virus have also been sequenced and are available in VirGen database (Kulkarni-Kale et al., 2004) of which 9 are vaccine strains and 12 are wild-type strains. The complete genome of the L-Zagreb mumps vaccine strain has been sequenced recently (Ivancic et al., 2005). Comparison of the nucleotide and protein sequences of L-Zagreb with other strains revealed that although the functional regions of HN, V and L proteins are conserved, there are substantial variations observed in the known antigenic regions (Ivancic et al., 2005). Furthermore, no molecular pattern was identified which can serve as a distinction marker between virulent and attenuated strains (Ivancic et al., 2005). The structures of HN of three viruses, viz., NDV, PIV and SV-5 that belong to the family *Paramyxoviridae* have been solved (Crennell et al., 2000; Takimoto et al., 2002; Lawrence et al., 2004; Yuan et al., 2005). The knowledge gained from these structures has been used to understand the mechanism of viral membrane fusion (Zaitsev et al., 2004; Lamb et al., 2006). HN has three functions, it recognizes the sialic acid containing receptors on the cell surface; promotes the fusion activity of F protein, viral penetration through the cell surface and removes the sialic acid from progeny virus particles to prevent self-agglutination (Crennell et al., 2000; Takimoto et al., 2002; Lamb et al., 2006). Due to its multi-functional nature, HN molecule is a target for development of antiviral drugs (Crennell et al., 2000). Similarly, due to its ability to elicit neutralizing antibodies, HN is a preferred candidate for vaccine development.

Sequence- and structure-based analyses of HN protein of mumps virus revealed that although there exists very high sequence similarity both at the nucleotide and protein level, various strains have acquired certain mutations which confer strain-specific properties. It is known that phenotypic properties such as antigenicity, immunogenicity and neurovirulence are the result of ‘spatio-temporal’ hierarchical processes. Exact mapping of the same at molecular level is difficult due to the fact that there exists complexity in terms of intermolecular interactions which may be discrete or continuous (Flower et al., 2003). Nevertheless, analyses of sequences and structures can help in providing the rationale to explain the phenotypic properties in related organisms.

Sequence-based analyses

The MSA revealed the positions of HN protein that are under selection pressure. There are about 23 and 44 ‘informative and variable’ sites respectively in vaccine and wild-type strains of mumps virus and both the groups share 15 such sites (see Table 2). Similarly a few ‘singleton variable’ sites are observed in both the groups. For example, T265A, S368R and R459P mutations are unique to SBL-1 and could contribute towards its strain specificity. From the plots of variability index, it is interesting to note that the wild-type strains have a large number of sites that are subjected to selection pressure and a wide variety of substitutions are observed among them. Further strain specificity is conferred by post-translational modifications. The attachment of the carbohydrate moiety is known to play an important role in the antigen–antibody interactions at molecular level (Petrescu et al., 2004) and variation of the immune response can be correlated to the extent of glycosylation (Lisowska, 2002; Palomo et al., 2000; Cole et al., 2004). The mutation at position 464 from N to K and H in the JL2 and JL5 can result in loss of potential glycosylation in these vaccine strains. Absence of carbohydrate moiety will alter the structural and chemical surface in these vaccine strains and contribute to its efficacy (Ong et al., 2005). Since this site is surrounded by the known neutralizing epitopes (see Fig. 2b), it can be hypothesized that the antibodies raised against these vaccines may not be able to neutralize the genotypes that are glycosylated at this position. Furthermore, this particular glycosylation site is flanked by mutations specific to vaccine strains viz., JL, JL-2, JL-5, Rubini and Sipar-02, indicating that this site may play a deterministic role in molecular recognition of antigen and offers a probable explanation for the observation that SBL-1 strain remains endemic in Europe (Orvell et al., 1997; Tecle et al., 1998; CDC, 2006).

MSA of 15 vaccine strains is shown as Annexure V. The variations observed have been marked by taking into account the amino acid predominant at the corresponding position in the MSA of wild-type strains. All the vaccine strains have a set of unique variations that may be responsible for attenuation. However, a few vaccine strains deviate minimally from the wild type while acquiring attenuation. For instance, out of the forty-seven mutations only four in L-Zagreb are such that the given amino acid is not the preferred amino acid in all wild-type strains. On the contrary, two Rubini vaccine strains show 18 and 19 such variations where the amino acid at the respective position is not the preferred one among the wild type (see Annexure V). As seen from Annexure V, the experimentally characterized epitope 329–340 is conserved among the vaccine strains. However, variations were observed in a few vaccine strains in epitope 265–288. Similarly, two sites vary in Rubini, JL-2 and JL-5 in the epitope 352–360. Apart from the above mentioned experimentally characterized epitopes, MSA also revealed the existence of a hypervariable region 462–474.

In addition to MSA of vaccine strains, pairwise alignments of 15 vaccine strains with 9 representative

Table 3

Sequence variation (100–%Similarity) between vaccine and wild-type strains computed using pairwise global alignments of HN protein of mumps virus

No.	Vaccine strain	%Difference	
		Min	Max
1	L-Zagrebvacc	0.9	1.9
2	L-Zagrebmas	0.9	1.9
3	Sipar02	0.3	2.4
4	Miyahara	0.9	2.1
5	Hoshino	0.5	2.2
6	Urabe	0.3	2.4
7	SkBv	0.0	2.4
8	Rubv	1.4	3.8
9	Rub1v	0.2	2.9
10	JL2	0.9	2.7
11	JL2m	0.9	2.7
12	JL2m1	1.4	3.4
13	JL5	1.7	3.3
14	JL5ma	1.7	3.3
15	JL5ma1	1.5	3.1

Note. representative wild-type strains of various genotypes as listed by Jin et al. (2005) have been used. For more details refer to Annexure V.

strains of genotypes A to E and G to J (as listed by Jin et al., 2005) were carried out. In the absence of HN sequence for Urb/Jap67 strain, 871005 (GenBank: AF314562) was selected to represent genotype B. No HN sequence was available for the representative strain or any other strain of genotype F. For genotypes K and L only partial sequence data for HN were available and hence were not included in the analysis. Global alignments were carried out by needle program of EMBOSS package (Rice et al., 2000). Annexure VI shows %identity, %similarity and %difference. %Difference is calculated as (100–%similarity). Table 3 shows the %difference between 15 vaccine strains and 10 representative sequences. Minimum deviation was observed between vaccine strain and the representative strain of the genotype to which it belongs. For example, SkBv, Rub1v, Sipar02 and Urabe deviate only by 0, 0.2, 0.3 and 0.3 respectively. The range of minimum and maximum deviation varies from 0.9–1.9 for L-Zagreb to 0.2–2.9 for Rub1v. These observations provide a rationale for the varying specificities of the vaccine strains for different genotypes.

Structure-based analyses

Comparison of 3D structures of four strains showed structural similarity with the HN of NDV indicating that the fold and the function of the mumps HN is conserved despite 65% sequence variation. Overall structural similarity was also observed among four strains and a few structural deviations were found only in the regions of mutations, as expected. It must be mentioned that none of the observed mutations are part of dimerization interface, although mutations at dimerization interface have been observed among members of the genus *Rubulavirus*. There are 24 mutations among the SBL-1 and three vaccine strains viz., L-Zagreb, JL2 and JL5 (Fig. 2b). Of these, 16 (259, 265, 266, 287, 288, 354, 356, 368, 372, 442, 462, 464, 468, 470, 473, 474) are present on one

surface (front view) of HN whereas only 8 (135, 161, 218, 279, 336, 459, 490) are present on the other surface (back view: not shown). Of the 24 mutations, twenty-three are accessible (%accessibility >25%) and only one at position 577 is buried in the core. Twelve of the sixteen mutations present on one surface are part of experimentally characterized epitopes. Furthermore, these 12 mutations are localized and predicted to be part of the conformational epitopes indicating that this region is a major site for antibody recognition as well as neutralization. Mutations are colored according to predominance of a given amino acid (Fig. 2b). Dark color indicates that the said amino acid is present in three of the four strains. The mutations, A265T, N266D, R368S, K468E, are shown in pink (SBL-1) and magenta (vaccine strains). Residues shown in magenta are identical in the vaccine strains viz., L-Zagreb, JL5, and JL2. The mutations, A265T and R368S, are specific only to SBL-1. All vaccine as well as wild-type strains except SBL-1 have T and S at the respective positions. Similarly, mutation N266D is specific to SBL-1 as all other wild-type strains and all other vaccine strains except Rubini have D at this position. In case of the mutation K468E, E is a preferred amino acid at this position in all wild-type and vaccine strains except SBL-1, Edinburgh-4/3 (wild type) and Rubini (vaccine strain). The fact that SBL-1 co-circulates in JL vaccinated population may be attributed to the mutations unique to SBL-1 (Orvell et al., 1997; Teclé et al., 1998).

Mutations, V259I, I287V, P354Q, E356D, S372N, V474A, are shown in blue (SBL-1, JL2, JL5) and cyan (L-Zagreb). Substitutions of V with I/A may not bring about a major change as they are known to be conformationally similar (Kolaskar and Kulkarni-Kale, 1992) and hence are favorable mutations. The mutation V259I is unique to L-Zagreb and may contribute towards the specificity of this vaccine strain. All wild-type and vaccine strains have V at this position. Similarly, I287V mutation is common at this position in all wild-type and vaccine strains. Majority of wild and vaccine strains except JL-2, JL-5 and Rubini are found to have V at this position. Similarly, the mutation E356D does not alter the biochemical property of the respective molecules. All wild-type strains except SBL-1, Kilham and Edinburgh-4/5 and all vaccine strains except JL-2, JL-5 and Rubini have E at this position. The mutation, P354Q is observed in JL-5 and L-Zagreb. All vaccine as well as wild-type strains have Q where as vaccine strains viz., JL-2 and Rubini and wild-type strains viz., SBL-1, Edinburgh-4/5 have P; Kilham strain has H and Zagreb cro69 strain has K at this position. Similarly S is a preferred residue at position 372 in the wild-type and vaccine strains including L-Zagreb. However, the wild-type strains viz., SBL-1, Edinburgh-4/5 and the vaccine strains viz., JL-2, JL-5, JL minor, major component have N (S372N). These analyses indicate that SBL-1 is very different from the vaccine strains that are in use. The vaccine strains JL-2, JL-5 and Rubini also have unique mutations (Annexure V).

The mutations V470I, N464K/H, S462L and T288K are shown in orange (SBL-1, JL5 and L-Zagreb) and yellow (JL2, JL5). All the above mutations are found in JL2 whereas JL5

has retained some residues of the wild type at a few positions but acquired mutations in others. The site N464 in SBL-1 is replaced by two different positively charged residues viz., K in JL2 and H in JL5 indicating that this site is under selection pressure in these vaccine strains. The vaccine strain of Smith-Kline Beecham live-attenuated Urabe AM9 vaccine (GenBank: AF314559) also has K at the equivalent position. L-Zagreb, like most of other vaccine and wild-type strains, is found to have N at this position (Annexure V). The wild-type strain, RW of genotype D (GenBank: M19933) has K at this position. Mutations T473I, S442Y are shown in dark green (SBL-1, L-Zagreb, JL2) and light green (JL5). Hence these mutations are unique to JL5 and may contribute to its strain specificity.

Mutations identified using MSA of vaccine strains (Annexure V) when mapped on the predicted 3D structure revealed that the residues of hypervariable region 462–474 are found to cluster on one of the surfaces of HN. Furthermore, the residues 462, 464, 468, 470, 473 and 474 are solvent accessible and are in close proximity to the known epitopes. This prompts us to propose that the region 462–474 is an additional antigenic site capable of playing a deterministic role in neutralization (see Fig. 2b).

Hence, both sequence- and structure-based analyses explain the fact that L-Zagreb has its own set of unique mutations acquired during the process of attenuation. However, it has been observed that the residues spanning known and predicted neutralizing epitopes are conserved in L-Zagreb with reference to wild-type strains, thereby rendering L-Zagreb the capacity to elicit strong and avid neutralizing antibodies. Thus, these studies explain the observed antigenic diversity and strain specificity of mumps virus apart from providing an insight into underlying molecular mechanisms responsible for observed variations in the efficacy of mumps vaccines.

Materials and methods

Compilation of data sets

Nucleotide and protein sequences of *HN* and *SH* genes of mumps virus were retrieved from GenBank (Benson et al., 2006) and GenPept (Wheeler et al., 2006) databases respectively. The 3D coordinates of orthologous proteins, which were used as templates in homology modeling, were obtained from PDB (Berman et al., 2000). The data on known antigenic regions and escape mutants were compiled from the published literature.

Data analyses

Pairwise global sequence alignments were carried out using the needle program of EMBOSS package (Rice et al., 2000). Multiple sequence alignments (MSA) were carried out using ClustalW (Chenna et al., 2003). Variability plots were obtained by a program developed in-house, which implements Wu and Kabat (1970) index. B cell epitopes of HN were predicted with the program ANTIGEN (Kolaskar and Tongaonkar, 1990) and

Conformational Epitope Prediction (CEP) Server (Kulkarni-Kale et al., 2005; Kolaskar and Kulkarni-Kale, 1999).

Molecular modeling

Three-dimensional structure of HN protein of mumps virus was predicted using knowledge-based homology modeling approach described previously (Kolaskar and Kulkarni-Kale, 1999; Kulkarni-Kale and Kolaskar, 2003). The Homology, Discover and Biopolymer modules of the InsightII suite of programs (version 2000) were used for this purpose (Accelrys Inc. USA).

The steps specific to the modeling of HN protein of mumps virus are as follows:

- Identification of strains for prediction of structure: four strains of mumps virus were identified for molecular modeling studies viz., one wild-type strain SBL-1 (genotype A; GenBank accession number: M55065) and three vaccine strains, namely L-Zagreb, JL5 and JL2 (sequenced at the Serum Institute of India, Pune, India and deposited in GenBank with accession numbers AY583323, AY584603 and AY584604 respectively).
- Identification of templates: the orthologues with known structures were identified using the program BLAST (Altschul et al., 1997) against PDB database. Crystal structures of HN protein of two members of *Paramyxoviridae* viz., human parainfluenza virus (PIV) PDB-ID: 1V3B, 1V3C, 1V3D, 1V3E and Newcastle Disease virus (NDV) PDB-ID: 1E8U, 1E8V, 1E8T, 1USR, 1USX were short-listed.
- Template evaluation and prioritization: the template structures were ranked according to resolution. A few structures are complexes of HN with a natural ligand and/or with small molecules. Templates were evaluated for multiple criteria such as geometry, stereochemistry and occupancy of dihedral angles in the allowed regions of Ramachandran plot. Structural alignment of templates (1E8U, 1E8V, 1E8T, 1USR, 1USX) was carried out to identify the structurally conserved and variable regions. The overall rmsd was found to be 0.3 Å. On the basis of the sequence similarity and the template evaluation studies, HN of NDV (PDB: 1E8U; Crennell et al., 2000; Takimoto et al., 2000) was chosen as a template.
- Identification of SCRs and loop regions: the structurally conserved regions (SCRs) and loops were identified using structural alignment of templates and multiple sequence alignment of HN protein of members of *Rubulavirus* genus. Care was taken not to allow breaks in the known secondary structural elements.
- Optimization of predicted structures: the amber all atom force field (Seibel et al., 1990) with distance-dependent dielectric constant was used to predict the 3D structures of HN protein. Optimization of structure was carried out using the Steepest Descents (SD) and Conjugate Gradient (CG) minimization methods till the average rms derivative criteria reached 0.01 and 0.001 kcal/mol/Å, respectively.
- Evaluation of models: the essential accuracy and correctness of the models were evaluated using various methods viz., PROSTAT (module in Homology), ProsaII (Sippl, 1993) and PROCHECK (Laskowski et al., 1993).

Acknowledgments

Perl scripts and useful comments by Shriram Bhosle are gratefully acknowledged. Dr. S.R.R. Reddy is profusely thanked for his help in improving the readability of the manuscript. Thanks to the editor and the anonymous reviewers for their constructive and enjoyable criticism. This work was supported by a contract as research grant to UKK by the Serum Institute of India Research Foundation. Molecular modeling facility funded by the Department of Biotechnology, Government of India at the Bioinformatics Centre, University of Pune, is acknowledged gratefully.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.virol.2006.09.040](https://doi.org/10.1016/j.virol.2006.09.040).

References

- Afzal, M.A., Pickford, A.R., Forsey, T., Heath, A.B., Minor, P.D., 1993. The Jeryl Lynn vaccine strain of mumps virus is a mixture of two distinct isolates. *J. Gen. Virol.* 74, 917–920.
- Afzal, M.A., Buchanan, J., Dias, J.A., Cordeiro, M., Bentley, M.L., Shorrock, C.A., Minor, P.D., 1997a. RT-PCR based diagnosis and molecular characterisation of mumps viruses derived from clinical specimens collected during the 1996 mumps outbreak in Portugal. *J. Med. Virol.* 52, 349–353.
- Afzal, M.A., Buchanan, J., Heath, A.B., Minor, P.D., 1997b. Clustering of mumps virus isolates by *SH* gene sequence only partially reflects geographical origin. *Arch. Virol.* 142, 227–238.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- Amexis, G., Fineschi, N., Chumakov, K., 2001. Correlation of genetic variability with safety of mumps vaccine Urabe AM9 strain. *Virology* 287, 234–241.
- Amexis, G., Rubin, S., Chizhikov, V., Pelloquin, F., Carbone, K., Chumakov, K., 2002. Sequence diversity of Jeryl Lynn strain of mumps virus: quantitative mutant analysis for vaccine quality control. *Virology* 300, 171–179.
- Apweiler, R., Hermjakob, H., Sharon, N., 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* 1473, 4–8.
- Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Wheeler, D.L., 2006. GenBank. *Nucleic Acids Res.* 34, D16–D20 (Database issue).
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. *Nucleic Acids Res.* 28, 235–242.
- Brown, E.G., Dimock, K., Wright, K.E., 1996. The Urabe AM9 mumps vaccine is a mixture of viruses differing at amino acid 335 of the hemagglutinin–neuraminidase gene with one form associated with disease. *J. Infect. Dis.* 174, 619–622.
- CDC, 2006. Update: multistate outbreak of mumps – United States, January 1 – May 2, 2006. *MMWR* 55, 559–563.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., Thompson, J.D., 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31, 3497–3500.

- Cole, K.S., Steckbeck, J.D., Rowles, J.L., Desrosiers, R.C., Montelaro, R.C., 2004. Removal of N-linked glycosylation sites in the V1 region of simian immunodeficiency virus gp120 results in redirection of B-cell responses to V3. *J. Virol.* 78, 1525–1539.
- Crennell, S., Takimoto, T., Portner, A., Taylor, G., 2000. Crystal structure of the multifunctional paramyxovirus hemagglutinin–neuraminidase. *Nat. Struct. Biol.* 7, 1068–1074.
- Cusi, M.G., Fischer, S., Sedlmeier, R., Valassina, M., Valensin, P.E., Donati, M., Neubert, W.J., 2001. Localization of a new neutralizing epitope on the mumps virus hemagglutinin–neuraminidase protein. *Virus Res.* 74, 133–137.
- Davies, D.R., Cohen, G.H., 1996. Interactions of protein antigens with antibodies. *Proc. Natl. Acad. Sci. U.S.A.* 93, 7–12.
- DeWeese-Scott, C., Moulton, J., 2004. Molecular modeling of protein function regions. *Proteins* 55, 942–961.
- Flower, D.R., McSparron, H., Blythe, M.J., Zygouri, C., Taylor, D., Guan, P., Wan, S., Coveney, P.V., Walshe, V., Borrow, P., Doytchinova, I.A., 2003. Computational vaccinology: quantitative approaches. *Novartis Found. Symp.* 254, 102–120.
- Flynn, M., Mahon, B.P., 2003. Cell mediated and humoral immune responses to mumps virus: recent developments. In: Pandalai, S.G. (Ed.), *Recent Research Developments in Virology*. Research Signpost, Trivandrum, India, pp. 97–115.
- Folb, P.I., Bernatowska, E., Chen, R., Clemens, J., Dodoo, A.N., Ellenberg, S.S., Farrington, C.P., John, T.J., Lambert, P.H., Macdonald, N.E., Miller, E., Salisbury, D., Schmitt, H.J., Siegrist, C.A., Wimalaratne, O., 2004. A global perspective on vaccine safety and public health: the Global Advisory Committee on Vaccine Safety. *Am. J. Public Health* 94, 1926–1931.
- Furesz, J., 2002. Safety of live mumps virus vaccines. *J. Med. Virol.* 67, 299–300.
- Galazka, A.M., Robertson, S.E., Kraigher, A., 1999. Mumps and mumps vaccine: a global review. *Bull. World Health Organ* 77, 3–14.
- Hobohm, U., Sander, C., 1994. Enlarged representation set of protein structures. *Protein Sci.* 3, 522–524.
- Ivancic, J., Gulija, T.K., Forcic, D., Baricevic, M., Jug, R., Mesko-Prejac, M., Mazuran, R., 2005. Genetic characterization of L-Zagreb mumps vaccine strain. *Virus Res.* 109, 95–105.
- Jin, L., Rima, B., Brown, D., Orvell, C., Tecele, T., Afzal, M., Uchida, K., Nakayama, T., Song, J.W., Kang, C., Rota, P.A., Xu, W., Featherstone, D., 2005. Proposal for genetic characterisation of wild-type mumps strains: preliminary standardisation of the nomenclature. *Arch. Virol.* 150, 1903–1909.
- Jorgensen, E.D., Collins, P.L., Lomedico, P.T., 1987. Cloning and nucleotide sequence of Newcastle disease virus hemagglutinin–neuraminidase mRNA: identification of a putative sialic acid binding site. *Virology* 156, 12–24.
- Kacprzak-Bergman, I., Zaleska, I., Jasek, J., 2001. Specific humoral immunity in mumps meningitis in children. *Med. Sci. Monit.* 7, 977–981.
- Kolaskar, A.S., Kulkarni-Kale, U., 1992. Sequence alignment approach to pick up conformationally similar protein fragments. *J. Mol. Biol.* 223, 1053–1061.
- Kolaskar, A.S., Kulkarni-Kale, U., 1999. Prediction of three-dimensional structure and mapping of conformational epitopes of envelope glycoprotein of Japanese encephalitis virus. *Virology* 261, 31–42.
- Kolaskar, A.S., Tongaonkar, P.C., 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.* 276, 172–174.
- Kovamees, J., Rydbeck, R., Orvell, C., Norrby, E., 1990. Hemagglutinin–neuraminidase (HN) amino acid alterations in neutralization escape mutants of Kilham mumps virus. *Virus Res.* 17, 119–129.
- Kulkarni-Kale, U., Kolaskar, A.S., 2003. Prediction of 3D structure of envelope glycoprotein of Sri Lanka strain of Japanese encephalitis virus. In: Chen, Yi-Ping Phoebe (Ed.), *Conferences in Research and Practice in Information Technology*, 19, pp. 87–96.
- Kulkarni-Kale, U., Bhosle, S., Manjari, G.S., Kolaskar, A.S., 2004. VirGen: a comprehensive viral genome resource. *Nucleic Acids Res.* 32, D289–D292.
- Kulkarni-Kale, U., Bhosle, S., Kolaskar, A.S., 2005. CEP: a conformational epitope prediction server. *Nucleic Acids Res.* 33, W168–W171.
- Lamb, R.A., Paterson, R.G., Jardetzky, T.S., 2006. Paramyxovirus membrane fusion: lessons from the F and HN atomic structures. *Virology* 344, 30–37.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. Procheck: a program to check the stereochemical quality of protein structure. *J. Appl. Crystallogr.* 26, 283–291.
- Lawrence, M.C., Borg, N.A., Streltsov, V.A., Pilling, P.A., Epa, V.C., Varghese, J.N., McKimm-Breschkin, J.L., Colman, P.M., 2004. Structure of the haemagglutinin–neuraminidase from human parainfluenza virus type III. *J. Mol. Biol.* 335, 1343–1357.
- Lim, C.S., Chan, K.P., Goh, K.T., Chow, V.T., 2003. Hemagglutinin–neuraminidase sequence and phylogenetic analyses of mumps virus isolates from a vaccinated population in Singapore. *J. Med. Virol.* 70, 287–292.
- Lisowska, E., 2002. The role of glycosylation in protein antigenic properties. *Cell. Mol. Life Sci.* 59, 445–455.
- Merz, D.C., Wolinsky, J.S., 1981. Biochemical features of mumps virus neuraminidases and their relationship with pathogenicity. *Virology* 114, 218–227.
- Mirza, A.M., Deng, R., Iorio, R.M., 1994. Site-directed mutagenesis of a conserved hexapeptide in the paramyxovirus hemagglutinin–neuraminidase glycoprotein: effects on antigenic structure and function. *J. Virol.* 68, 5093–5099.
- Nagai, T., Okafuji, T., Miyazaki, C., Ito, Y., Kamada, M., Kumagai, T., Yuri, K., Sakiyama, H., Miyata, A., Ihara, T., Ochiai, H., Shimomura, K., Suzuki, E., Torigoe, S., Igarashi, M., Kase, T., Okuno, Y., Nakayama, T., 2006. A comparative study of the incidence of aseptic meningitis in symptomatic natural mumps patients and monovalent mumps vaccine recipients in Japan. *Vaccine* [Jan 31, Epub ahead of print].
- Ong, G., Goh, K.T., Ma, S., Chew, S.K., 2005. Comparative efficacy of Rubini, Jeryl-Lynn and Urabe mumps vaccine in an Asian population. *J. Infect.* 51, 294–298.
- Orvell, C., Alsheikhly, A.R., Kalantari, M., Johansson, B., 1997. Characterization of genotype-specific epitopes of the HN protein of mumps virus. *J. Gen. Virol.* 78, 3187–3193.
- Palomo, C., Cane, P.A., Melero, J.A., 2000. Evaluation of the antibody specificities of human convalescent-phase sera against the attachment (G) protein of human respiratory syncytial virus: influence of strain variation and carbohydrate side chains. *J. Med. Virol.* 60, 468–474.
- Paterson, R.G., Lamb, R.A., 1990. RNA editing by G-nucleotide insertion in mumps virus P-gene mRNA transcripts. *J. Virol.* 64, 4137–4145.
- Petrescu, A.J., Milac, A.L., Petrescu, S.M., Dwek, R.A., Wormald, M.R., 2004. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* 14, 103–114.
- Pipkin, P.A., Afzal, M.A., Heath, A.B., Minor, P.D., 1999. Assay of humoral immunity to mumps virus. *J. Virol. Methods* 79, 219–225.
- Rafiefard, F., Johansson, B., Tecele, T., Orvell, C., 2005. Characterization of mumps virus strains with varying neurovirulence. *Scand. J. Infect. Dis.* 37, 330–337.
- Ramachandran, G.N., Sasisekharan, V., 1968. Conformation of polypeptides and proteins. *Adv. Protein Chem.* 23, 283–438.
- Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16, 276–277.
- Rima, B.K., Alexander, D.J., Billeter, M.A., Collins, P.L., Kingsbury, D.W., Lipkind, M.A., Nagai, Y., Orvell, C., Pringle, C.R., ter Meulen, V., 1995. Paramyxoviridae. In: Murphy, F.A., Fauquet, C.M., Bishop, D.H.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., Summers, M.D. (Eds.), *Virus Taxonomy*. Sixth Report of the International Committee of Viruses. Springer-Verlag, New York, pp. 268–274.
- Rost, B., O'Donoghue, S., 1997. Sisyphus and prediction of protein structure. *Comput. Appl. Biosci.* 13, 345–356.
- Rubin, S.A., Pletnikov, M., Carbone, K.M., 1998. Comparison of the neurovirulence of a vaccine and a wild-type mumps virus strain in the developing rat brain. *J. Virol.* 72, 8037–8042.
- Rubin, S.A., Pletnikov, M., Taffs, R., Snoy, P.J., Kobasa, D., Brown, E.G., Wright, K.E., Carbone, K.M., 2000. Evaluation of a neonatal rat model for prediction of mumps virus neurovirulence in humans. *J. Virol.* 74, 5382–5384.
- Rubin, S.A., Afzal, M.A., Powell, C.L., Bentley, M.L., Auda, G.R., Taffs, R.E., Carbone, K.M., 2005. The rat-based neurovirulence safety test for the assessment of mumps virus neurovirulence in humans: an international collaborative study. *J. Infect. Dis.* 191, 1123–1128.

- Saika, S., Kidokoro, M., Aoki, A., Ohkawa, T., 2004. Neurovirulence of mumps virus: intraspinal inoculation test in marmosets. *Biologicals* 32, 147–152.
- Saito, H., Takahashi, Y., Harata, S., Tanaka, K., Sano, T., Suto, T., Yamada, A., Yamazaki, S., Morita, M., 1996. Isolation and characterization of mumps virus strains in a mumps outbreak with a high incidence of aseptic meningitis. *Microbiol. Immunol.* 40, 271–275.
- Sauder, C.J., Vandenburgh, K.M., Iskow, R.C., Malik, T., Carbone, K.M., Rubin, S.A., 2006. Changes in mumps virus neurovirulence phenotype associated with quasispecies heterogeneity. *Virology* 350, 48–57.
- Seibel, G., Singh, U.C., Weiner, P.K., Caldwell, J., Kollman, P., 1990. AMfBER 3.0 revision A. University of California at San Francisco, San Francisco.
- Sippl, M.J., 1993. Recognition of errors in three-dimensional structures of proteins. *Proteins* 17, 355–362.
- Smith-Gill, S.J., 1996. Molecular recognition of lysozyme by monoclonal antibodies. *EXS* 75, 277–300.
- Takahashi, M., Nakayama, T., Kashiwagi, Y., Takami, T., Sonoda, S., Yamanaka, T., Ochiai, H., Ihara, T., Tajima, T., 2000. Single genotype of measles virus is dominant whereas several genotypes of mumps virus are co-circulating. *J. Med. Virol.* 62, 278–285.
- Takimoto, T., Taylor, G.L., Crennell, S.J., Scroggs, R.A., Portner, A., 2000. Crystallization of Newcastle disease virus hemagglutinin–neuraminidase glycoprotein. *Virology* 270, 208–214.
- Takimoto, T., Taylor, G.L., Connaris, H.C., Crennell, S.J., Portner, A., 2002. Role of the hemagglutinin–neuraminidase protein in the mechanism of paramyxovirus–cell membrane fusion. *J. Virol.* 76, 13028–13033.
- Tanaka, K., Baba, K., Okada, S., Okuno, Y., Yamanishi, K., Ueda, S., Takahashi, M., Yamada, A., 1992. Nasal antibody response to mumps virus after vaccination and natural infection. *Vaccine* 10, 824–827.
- Tecle, T., Johansson, B., Jecic, A., Forsgren, M., Orvell, C., 1998. Characterization of three co-circulating genotypes of the small hydrophobic protein gene of mumps virus. *J. Gen. Virol.* 79, 2929–2937.
- Tecle, T., Bottiger, B., Orvell, C., Johansson, B., 2001. Characterization of two decades of temporal co-circulation of four mumps virus genotypes in Denmark: identification of a new genotype. *J. Gen. Virol.* 82, 2675–2680.
- Tecle, T., Mickiene, A., Johansson, B., Lindquist, L., Orvell, C., 2002. Molecular characterisation of two mumps virus genotypes circulating during an epidemic in Lithuania from 1998 to 2000. *Arch. Virol.* 147, 243–253.
- Utz, S., Richard, J.L., Capaul, S., Matter, H.C., Hrisoho, M.G., Muhlemann, K., 2004. Phylogenetic analysis of clinical mumps virus isolates from vaccinated and non-vaccinated patients with mumps during an outbreak, Switzerland 1998–2000. *J. Med. Virol.* 73, 91–96.
- Van Regenmortel, M.H., 1998. Mimotopes, continuous paratopes and hydrophobic complementarity: novel approximations in the description of immunochemical specificity. *J. Dispers. Sci. Technol.* 19, 1199–1219.
- Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvermin, V., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S., Geer, L.Y., Helms, W., Kapustin, Y., Kenton, D.L., Khovayko, O., Lipman, D.J., Madden, T.L., Maglott, D.R., Ostell, J., Pruitt, K.D., Schuler, G.D., Schriml, L.M., Sequeira, E., Sherry, S.T., Sirotkin, K., Souvorov, A., Starchenko, G., Suzek, T.O., Tatusov, R., Tatusova, T.A., Wagner, L., Yaschenko, E., 2006. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 34, D173–D180.
- Wolinsky, J.S., 1996. *Mumps Virus*. Lippincott-Raven, Philadelphia.
- Wu, T.T., Kabat, E.A., 1970. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* 132, 211–250.
- Wu, L., Bai, Z., Li, Y., Rima, B.K., Afzal, M.A., 1998. Wild type mumps viruses circulating in China establish a new genotype. *Vaccine* 16, 281–285.
- Yates, P.J., Afzal, M.A., Minor, P.D., 1996. Antigenic and genetic variation of the HN protein of mumps virus strains. *J. Gen. Virol.* 77, 2491–2497.
- Yuan, P., Thompson, T.B., Wurzburg, B.A., Paterson, R.G., Lamb, R.A., Jardetzky, T.S., 2005. Structural studies of the parainfluenza virus 5 hemagglutinin–neuraminidase tetramer in complex with its receptor, sialyllactose. *Structure* 13, 803–815.
- Zaitsev, V., von Itzstein, M., Groves, D., Kiefel, M., Takimoto, T., Portner, A., Taylor, G., 2004. Second sialic acid binding site in Newcastle disease virus hemagglutinin–neuraminidase: implications for fusion. *J. Virol.* 78, 3733–3741.